Optimal Control Approach to Image Registration

Stephen Taiwo Salako

Department of Mathematics and Science, Our Lady of the Lake University, United States

Copyright©2018 by authors, all rights reserved. Authors agree that this article remains permanently open access under the terms of the Creative Commons Attribution License 4.0 International License

Abstract We discuss the orthogonality problem of moving grids, where we minimize a cost function with a regularization term and improve the orthogonality of grids by making the angle between grids close to 90°. The initial grid used is obtained by a well-established method known as the grid deformation method. We will then replace the cost function in the orthogonality problem with sum of squared differences (SSD) to discuss the image registration problem. We will discuss the non-uniqueness of solutions, existence of optimal solutions and prove the existence of Lagrange multipliers of the image registration problem using the Direct Method in Calculus of Variation and then derive an optimality system based on the construction of a Lagrangian functional from which optimal transformations can be calculated.

Keywords Image registration, Lagrange Multipliers, Orthogonality, Grid Deformation Method

1. Introduction

In this paper we use the moving grid method. The moving grid method forms the foundation for the optimization of grid orthogonality as well as the image registration problem.

1.2. Grid Deformation Method

The grid deformation method is used for the construction of differentiable and invertible transformations to solve mesh adaptation problems. To solve a differential equation numerically by finite difference, finite element or finite volume methods, a good grid is needed for discretization of the physical domain.

The idea of this method is to move nodes with correct velocities so that the nodal mapping has a desirable determinant.

There are numerous approaches of grid generation as discussed in [8]. In this paper we use the moving grid method.

Harmonic maps are widely used to generate a body-fitted structured grid on a general domain.

In 2D, due to the Rado’s Theorem which informally states that any “nice looking” shape without holes can be smoothly deformed into a disk.

Theorem 1.1 [Rado’s Theorem]: Suppose Ω is an open, connected and convex subset of the Euclidean space R² with smooth boundary ∂Ω and let D denote the unit disk. Then, given any homeomorphism μ : ∂D → ∂Ω, there exists a unique harmonic function u : D → Ω such that u = μ on ∂D and u is a diffeomorphism.

The 3D version of Rado’s theorem is false. The grids may fold in theory (and in practice).

The not so stable version of the grid deformation method was developed in [1], [2], [3]; it was improved in [4] and used with a finite-volume solver in flow calculations in [5]. A 2D version of the method was proposed in [6] and used with a discontinuous Galerkin finite-element method in solving a convection diffusion problem in [7].

The grid deformation method origins are based on a deformation scheme used by J. Moser [8] in differential geometry. The following result is taken from B. Dacorogna and J. Moser [9]:

1.2 Image Registration

An image is a 2 dimensional array of pixels with assigned brightness values. A pixel is the smallest item of information in an image. Each pixel is a portion of the original image. With more portions we get a better representation of the image.

Mathematically in [12], an image is defined as a d-dimensional function \( b : \mathbb{R}^d \rightarrow \mathbb{R} \) if

1. \( b \) is compactly supported
2. \( 0 \leq b(x) < \infty \) for all \( x \in \mathbb{R}^d \) and \( d \in \mathbb{N} \)
3. \( \int_{\mathbb{R}^d} b(x) \, dx \) is finite, for \( k > 0 \).

Further details can be found in [12].

Image registration is the process of overlaying two images of the same scene taken at different times, from different viewpoints; or images of different, but related scenes. It could also be from the same or different imaging modalities.

Some popular imaging modalities are x-ray, computed
Optimal Control Approach to Image Registration

axial tomography (CAT), magnetic resonance imaging (MRI), positron emission tomography (PET), functional MRI (fMRI), ultrasound (US).

Image registration is the process of finding an optimal geometric transformation between corresponding imaging data. In practice, the concrete type of the geometric transformation as well as the notions of optimal and corresponding depends on the specific application.

The image registration problem can be phrased in only a few words: given a Reference $R$ and Template image $T$, $R, T : \mathbb{R}^d \rightarrow \mathbb{R}$ (smooth, compact support), find a suitable transformation $\Phi : \mathbb{R}^d \rightarrow \mathbb{R}^d$ such that the transformed template becomes similar to the reference, i.e. $T \circ \Phi \approx R$.

Notice that the reference image has undergone a slight rotation. Our goal is to get the Template image to be as similar as possible to the Reference image by means of a transformation. The two images in Figure 1.1 are provided using the same imaging modality i.e. MRI.

Image registration is a problem often encountered in many applications areas like, for example, astro- and geophysics, computer vision and medicine. For an overview, see, [13], [14], [15] and [16], Brown (1992), Maintz & Viergever (1998), Maurer & Fitzpatrick (1993), van den Elsen et al (1993).

Image registration is an ill-posed problem, small changes to the input images can lead to completely different registration results.

As a result of this solutions are not unique.

For example,

Suppose we want to register the reference and template images above for simplicity reasons allowing only rigid transformation i.e. rotations and translations.

We find several solutions namely a pure translation, a rotation of 180°, a rotation of 90° followed by a translation and so on [12].

We can divide the applications of image registration into four main groups according to the manner of the image acquisition as follows:

1. Different viewpoints (multiview analysis): The goal is to gain a larger 2D view or a 3D representation of the scanned scene
(2) Different times (multitemporal analysis): Motion tracking, medical imaging monitoring of the healing therapy, monitoring of the tumor evolution
(3) Computer vision shape
(4) Different sensors (multimodal analysis)

Details about Computer vision shape and Different sensors can be found in [17].
In regards to medical images, registration can be achieved as mentioned earlier by using different imaging modalities based on specific needs. Using the same modality for a patient, monitoring and qualifying disease progress over time can be done. If tissue analysis is required, deformation monitoring can be done. Now, if different modalities are used with a patient, correction for different patient position between scans and link between structural and functional images can be done.

2. Optimization of Grid Orthogonality

In this chapter we will discuss about optimizing grid orthogonality by minimizing a cost function involving the orthogonal functional and penalty term. The goal is to keep the size distribution of the grids.

We will discuss what has been done before and the significant improvement we have made in this thesis. Analysis of the problem will be shown and numerical results based on the analysis will be presented in this thesis.

Let us first discuss briefly about numerical generation of grids.

2.1. Numerical Grid Generation

Numerical grid generation has now become a fairly common tool for use in the numerical solution of partial differential equations on arbitrarily shaped regions.

Numerical grid generation can be thought of as a procedure for the orderly distribution of observers, or sampling stations over a physical field in such a way that efficient communication among the observers is possible and that all physical phenomena in the entire continuous field may be represented with sufficient accuracy by this finite collection of observations. Another way to think of the grid is as the structure on which the numerical solution is built.

The numerical solution of partial differential equations requires some discretization of the field into a collection of points or elemental volumes.

The differential equations are approximated by a set of algebraic equations on this collection, and this system of algebraic equations is then solved to produce a set of discrete values which approximates the solution of the partial differential system over the field. General methods of grid generation as well as what is discussed above can be found in [18].

Note: Algorithms have been produced lately to solve ODEs and PDEs by moving grids, the latest papers which show its effectiveness can be read in [19], [20], [21], [22], [23], [24], [32] and [33]. These references provide the technique that show the moving mesh algorithm and the discrete problem can be patched up to find the solution on adaptive mesh. In addition these papers also provide methods where the problems are nonlinear and Jacobian needs to be used to solve these problems. These adaptive procedures are efficient and do not lead to ill conditioned matrix as low order polynomial interpolations are sufficient.

2.1.1. Variational Method of Improving Grid Generation

Properties of grids such as length of grid lines or smoothness, area or volume of cells, and orthogonality of grid lines are controlled by the minimization of a functional [25].

Brackbill and Saltzman in [26] combined these functionals into one functional, which is given as:

\[ I = \lambda_s I_s + \lambda_v I_v + \lambda_o I_o \]

where,
\[ \lambda_s, \lambda_v, \lambda_o \] are non-negative numbers.

It is shown by calculation in [26] that numerical solutions which minimize \( I \) is obtained for finite values of \( \lambda_o, \lambda_v \).

\( \lambda_o \) is a non-negative number that is associated to the volume weighted measure, \( I_o \).

Guojun Liao in [25] followed the approach of Roache and Steinberg in 1991 but used a different smoothness functional \( I_s \).

In [27], the functionals are:

\[ I_s = \frac{1}{2} \int_{\Omega} \left( |\nabla x|^2 + |\nabla y|^2 \right) d\zeta d\phi, \]
\[ I_v = \frac{1}{2} \int_{\Omega} \left( \frac{D(x,y)}{D(\zeta,\phi)} \right) d\zeta d\phi, \]
\[ I_o = \frac{1}{2} \int_{\Omega} \left( x_\zeta x_\zeta + y_\zeta y_\zeta \right) d\zeta d\phi \]

Where,
\( \nabla x \) = the gradient of \( x(\zeta, \phi) \) and \( \frac{D(x,y)}{D(\zeta,\phi)} \) is the Jacobian determinant.

In [28], it was shown that the volume is the most important property to control, followed by smoothness and orthogonality.

When the volume control is given more weight than smoothness, with a little orthogonality control, this technique seems to work best. It was observed that none of the grids that were generated with a significant amount of volume folded. The concern of grid folding in We enforce the Jacobian determinant of the grid generated to be strictly positive through the control of a monitor function \( f \).
A simpler model was studied in [25]. This model involved only a weighted sum of the smoothness control functional, $I_s$, and the volume (area) control functional, $I_v$, i.e.

$$ I = \alpha^2 I_s + \beta^2 I_v $$

The existence of a minimum of $I$ in a Sobolev space is proven by the direct method in calculus of variations. In [25], an explanation is given as to why $I_v$ is used to control cell volume.

Since in the theory of calculus of variation, convexity is required, a weak sensed convexity is proven i.e. poly convexity.

This guarantees that the Euler-Lagrange equation is elliptic in the sense of Agmon, Douglis and Nirenberg which is also referred to as the Legendre-Hadamard condition.

J.U. Brackbill et al in [26] quoted [29] in that their $I_v$ and $I_o$ (volume weighted measure) cannot be minimized separately because they do not have unique solutions, an example is used to prove this.

$I_s$ is shown to have a unique solution in [30], so to formulate a minimization problem with unique solutions as in the penalty method, $I$ is minimized in the manner below:

$$ I = I_s + \lambda_v I_v + \lambda_o I_o $$

Numeric solutions which minimize $I_s$ are obtained for finite values of $\lambda_o$ and $\lambda_v$.

Liao in [31] proved that there exists a $C^{0, \alpha}$ regularity of minimum of a functional related to the grid-generation problem.

The functional discussed is a linear combination of the length functional, the volume functional and a higher order term.

The inclusion of the latter guarantees the existence and regularity of a minimum via the direct method for calculus of variations.

In this paper, we minimize $I$ with $I_s$, the smoothness functional without the volume or area functional, $I_v$, i.e. we minimize the orthogonal functional with a penalty term

$$ \frac{\lambda}{2} \int_D \| \nabla \phi \|^2 $$

to make the problem more elliptic.

We then find the Euler-Lagrange equations and solve for $u$ numerically, which gives us the orthogonal grids.

2.1.2. Formulation of the Problem

Minimize: $\int_{D} (\phi_{x_i} \cdot \phi_{x_j})^2$

Let us find the Euler-Lagrange equations that the minimizing function must satisfy.

Let us write the functional as:

$$ I = \frac{1}{2} \int_D \left( \phi_{x_i} \phi_{x_j} + \mu (\phi_{x_k} \phi_{x_k} + \phi_{x_l} \phi_{x_l}) \right)^2 + \frac{\lambda}{2} \int_D \| \nabla \phi \|^2 $$

Where:

$$ \phi = (\phi_1, \phi_2) $$

Now, let us introduce a small variation of $\phi$; namely,

$$ \phi \rightarrow \phi + \mu \delta \phi $$

Now,

$$ I_{\phi} = 0 $$

$$ \frac{\partial I(\phi + \mu \delta \phi)}{\partial \phi_i} = \mu \delta \phi_i $$

$$ \frac{\partial I}{\partial \phi_i} = \frac{1}{2} \int_D \left( \phi_{x_i} \phi_{x_j} + \mu (\phi_{x_k} \phi_{x_k} + \phi_{x_l} \phi_{x_l}) \right)^2 + \frac{\lambda}{2} \int_D \| \nabla \phi \|^2 $$

$$ \frac{\partial I}{\partial \phi_i} = \mu \delta \phi_i + \mu \delta \phi_{x_j} \phi_{x_j} + \mu \delta \phi_{x_k} \phi_{x_k} + \mu \delta \phi_{x_l} \phi_{x_l} $$

$$ \frac{\partial I}{\partial \phi_i} = \mu \delta \phi_i + \mu \delta \phi_{x_k} \phi_{x_k} + \mu \delta \phi_{x_l} \phi_{x_l} $$

$$ \frac{\partial I}{\partial \phi_i} = \mu \delta \phi_i + \mu \delta \phi_{x_k} \phi_{x_k} + \mu \delta \phi_{x_l} \phi_{x_l} $$

Figure 2.1. Illustration of orthogonality goal
\[\Delta \phi = - \text{div}(A, B)/\lambda \] (2.1.2.9)

\[\Delta \phi = - \text{div}(C, D)/\lambda \] (2.1.2.10)

Let,

\[\phi(x) = \phi_o(x) + u(x)\]

\(u(x)\) is a displacement to be added to the initial grid \(\phi_o(x)\).

The initial grid \(\phi_o(x)\) is generated by the deformation method version 3 discussed in chapter 1.

We use a \(D = 25 \times 25\) grid, \(\mathbf{n} \cdot \mathbf{u} = 0\) on \(\partial D\). Now,

Substitute \(\phi(x) = \phi_o(x) + u(x)\) into (2.1.2.9) and (2.1.2.10) respectively, to get:

\[\Delta \phi(x) = \Delta \phi_o(x) + \Delta u(x) = - \text{div}(A, B)/\lambda\] (2.1.2.11)

\[\Delta \phi(x) = \Delta \phi_o(x) + \Delta u_2(x) = - \text{div}(C, D)/\lambda\] (2.1.2.12)

We find \(u_1\) and \(u_2\) from:
\[ \Delta u_1(x) = \frac{-\text{div}(A, B)}{\lambda} - \Delta \phi_1(x) \]  
(2.1.13)

\[ \Delta u_2(x) = \frac{-\text{div}(C, D)}{\lambda} - \Delta \phi_2(x) \]  
(2.1.14)

\[ \Delta u_1(x) = \frac{-A_{x_1} - B_{x_1}}{\lambda} - \Delta \phi_1(x) \]  
(2.1.15)

\[ \Delta u_2(x) = -\frac{C_{x_1} - D_{x_1}}{\lambda} - \Delta \phi_2(x) \]  
(2.1.16)

\[ \Delta u_1(x) = \frac{-\left(\phi_{i_1, x_{i_1}}, \phi_{i_1, x_{21}}, \phi_{i_1, x_{22}}\right) - \left(\phi_{i_1, x_{i_1}}, \phi_{i_1, x_{21}}, \phi_{i_1, x_{22}}\right)}{\lambda} - \Delta \phi_1(x) \]  
(2.1.17)

\[ \Delta u_2(x) = \frac{-\left(\phi_{i_2, x_{i_2}}, \phi_{i_2, x_{21}}, \phi_{i_2, x_{22}}\right) - \left(\phi_{i_2, x_{i_2}}, \phi_{i_2, x_{21}}, \phi_{i_2, x_{22}}\right)}{\lambda} - \Delta \phi_2(x) \]  
(2.1.18)

2.2. Numerical Scheme

We obtain the initial grid \( \varphi_0 \), by the deformation method, discussed in chapter 1. When we solve the Euler-Lagrange equations (2.1.17) and (2.1.18) we get the deformation of grids \( u \). The composition grids \( \varphi_0 + u \) give us the orthogonal grids i.e. every point is moved a little bit. We use the Neumann boundary condition to be \( 0 \). i.e. \( n \cdot u = 0 \). Computation of \( \varphi_0 \) was done by calling a Laplace solver (subroutine). The initial grid \( \varphi_0 \) is only needed once. The field variables \( A, B, C, D \) in (2.1.15) and (2.1.16) are computed, these are the derivatives of \( \varphi_0 + u \). To solve the right hand side of (2.1.15) and (2.1.16), we call a poisson solver. When solving SOR we introduce a relaxation factor (real \( \text{vomega} \)) and tolerance (tolr).

There is a control about the poisson solver: bdfactor. If bdfactor = 0, the computation is not conducted on the boundary but if the bdfactor = 1, the computation is conducted on the boundary.

Tolerance is needed to help satisfy (2.1.15) and (2.1.16) i.e. to show that our equation is correct. The computational residue gets closer to the tolerance, 0.001 is sufficient.

Two kinds of boundaries are used: the fixed boundary and the mirror boundary, because we don’t know the correct boundary.

2.2.1. Dirichlet Boundary Condition

Complete computation on boundary is conducted, using the one-sided difference scheme.

2.2.2. Neumann Boundary Condition

Imposes the imaginary grid points \( \varphi_x \) and the variation \( u \) near the boundary according to symmetric relation.

2.3. Numerical Experimentation

This chapter presents several numerical results showing the improvement in orthogonality of the grid by first producing an initial grid \( \varphi_0 \) by the deformation method and then composing with \( u \), which is found by solving the Euler-Lagrange Equations, and then the composition \( \varphi_0 + u \) produces the orthogonal grid.

The initial grid \( \varphi_0 \) is a 25 x 25 grid.

![Initial Grid, \( \varphi_0 \)](image)

CASE 1: The controlling parameters are as follows:

| Variables   | Value | Explanation                                      |
|-------------|-------|--------------------------------------------------|
| \( \lambda \) | 1     | The control parameter in the new method          |
| rsdIMax     | 0.001 | The tolerating value for the iteration in the new method |
| iterMax     | 200   | The maximum number for the iteration in the new method |
| vomega      | 1     | The control parameter for solving SOR            |
| tolr        | 0.001 | The tolerating value for the iteration in SOR     |
| iBDtype     | 1     | boundary type:1-dirichlet boundary; 2-neumann boundary |
| bdfactor    | 0     | bdfactor: 1-do computation on boundary, 0-not do  |
| \( \mu \)   | 0.2   | the quantity level for adding the change u       |
CASE 2: The controlling parameters are as follows:

**Table 2.2.** \( \lambda = 0.9, \mu = 0.2 \)

| Variables  | Value | Explanation                                      |
|------------|-------|-------------------------------------------------|
| \( \lambda \) | 0.9   | The control parameter in the new method          |
| rsdlMax    | 0.001 | The tolerating value for the iteration in the new method |
| iterMax    | 200   | The maximum number for the iteration in the new method |
| vomega     | 1     | The control parameter for solving SOR            |
| tolr       | 0.001 | The tolerating value for the iteration in SOR    |
| iBDtype    | 2     | boundary type:1-dirichlet boundary, 2-neumann boundary |
| bdfactor   | 0     | bdfactor: 1-do computation on boundary, 0-not do |
| \( \mu \)  | 0.2   | the quantity level for adding the change u       |

CASE 3: The controlling parameters are as follows:

**Table 2.3.** \( \lambda = 0.7, \mu = 0.4 \)

| Variables  | Value | Explanation                                      |
|------------|-------|-------------------------------------------------|
| \( \lambda \) | 0.7   | The control parameter in the new method          |
| rsdlMax    | 0.001 | The tolerating value for the iteration in the new method |
| iterMax    | 200   | The maximum number for the iteration in the new method |
| vomega     | 1     | The control parameter for solving SOR            |
| tolr       | 0.001 | The tolerating value for the iteration in SOR    |
| iBDtype    | 2     | boundary type:1-dirichlet boundary, 2-neumann boundary |
| bdfactor   | 0     | bdfactor: 1-do computation on boundary, 0-not do |
| \( \mu \)  | 0.4   | the quantity level for adding the change u       |
CASE 4: The controlling parameters are as follows:

| Variables | Value | Explanation |
|-----------|-------|-------------|
| $\lambda$ | 0.7   | The control parameter in the new method |
| rsdMax    | 0.001 | The tolerating value for the iteration in the new method |
| iterMax   | 200   | The maximum number for the iteration in the new method |
| vomega    | 1     | The control parameter for solving SOR |
| tol       | 0.001 | The tolerating value for the iteration in SOR |
| iBType    | 2     | boundary type: 1-dirichlet boundary; 2-neumann boundary |
| bdfactor  | 0     | bdfactor: 1-do computation on boundary, 0-not do |
| $\mu$     | 0.9   | the quantity level for adding the change $u$ |

CASE 5: The controlling parameters are as follows:

| Variables | Value | Explanation |
|-----------|-------|-------------|
| $\lambda$ | 0.69  | The control parameter in the new method |
| rsdMax    | 0.001 | The tolerating value for the iteration in the new method |
| iterMax   | 200   | The maximum number for the iteration in the new method |
| vomega    | 1     | The control parameter for solving SOR |
| tol       | 0.001 | The tolerating value for the iteration in SOR |
| iBType    | 2     | boundary type: 1-dirichlet boundary; 2-neumann boundary |
| bdfactor  | 0     | bdfactor: 1-do computation on boundary, 0-not do |
| $\mu$     | 0.4   | the quantity level for adding the change $u$ |
CASE 6: The controlling parameters are as follows:

Table 2.6. \( \lambda = 0.69, \ \mu = 0.9 \)

| Variables | Value | Explanation |
|-----------|-------|-------------|
| \( \lambda \) | 0.69  | The control parameter in the new method |
| rsdIMAX   | 0.001 | The tolerating value for the iteration in the new method |
| iterMax   | 200   | The maximum number for the iteration in the new method |
| vomega    | 1     | The control parameter for solving SOR |
| tolr      | 0.001 | The tolerating value for the iteration in SOR |
| iBdtype   | 2     | boundary type: 1-dirichlet boundary, 2-neumann boundary |
| bdfactor  | 0     | bdfactor: 1-do computation on boundary, 0-not do |
| \( \mu \) | 0.9   | the quantity level for adding the change \( u \) |

CASE 7: The controlling parameters are as follows:

Table 2.7. \( \lambda = 0.75, \ \mu = 0.4 \)

| Variables | Value | Explanation |
|-----------|-------|-------------|
| \( \lambda \) | 0.75  | The control parameter in the new method |
| rsdIMAX   | 0.001 | The tolerating value for the iteration in the new method |
| iterMax   | 200   | The maximum number for the iteration in the new method |
| vomega    | 1     | The control parameter for solving SOR |
| tolr      | 0.001 | The tolerating value for the iteration in SOR |
| iBdtype   | 2     | boundary type: 1-dirichlet boundary, 2-neumann boundary |
| bdfactor  | 0     | bdfactor: 1-do computation on boundary, 0-not do |
| \( \mu \) | 0.4   | the quantity level for adding the change \( u \) |
CASE 8: The controlling parameters are as follows:

| Variables | Value | Explanation |
|-----------|-------|-------------|
| $\lambda$ | 50    | The control parameter in the new method |
| rsdIMax   | 0.001 | The tolerating value for the iteration in the new method |
| iterMax   | 200   | The maximum number for the iteration in the new method |
| vomega    | 1     | The control parameter for solving SOR |
| tolr      | 0.001 | The tolerating value for the iteration in SOR |
| iBDtype   | 2     | boundary type: 1-dichlet boundary; 2-neumann boundary |
| bdfactor  | 0     | bdfactor: 1-do computation on boundary, 0-not do |
| $\mu$     | 0.2   | the quantity level for adding the change $u$ |

CASE 9: The controlling parameters are as follows:

| Variables | Value | Explanation |
|-----------|-------|-------------|
| $\lambda$ | 100   | The control parameter in the new method |
| rsdIMax   | 0.001 | The tolerating value for the iteration in the new method |
| iterMax   | 200   | The maximum number for the iteration in the new method |
| vomega    | 1     | The control parameter for solving SOR |
| tolr      | 0.001 | The tolerating value for the iteration in SOR |
| iBDtype   | 2     | boundary type: 1-dichlet boundary; 2-neumann boundary |
| bdfactor  | 0     | bdfactor: 1-do computation on boundary, 0-not do |
| $\mu$     | 0.2   | the quantity level for adding the change $u$ |
2.4. Discussion

In 2.1.2.9 and 2.1.2.10, $\lambda$ was found to work if between 0.6-1.0, otherwise the computation is unstable.

In case 1, with $\lambda = 1$ the computation converges after 7 iteration. Slight orthogonality is noticed [Figure 2.3].

In case 2 with $\lambda = 0.9$, the computation converges after 6 iterations, more orthogonality is noticed. [Figure 2.4]

In case 3 with $\lambda = 0.7$, the computation converges after 11 iterations, significant orthogonal grid noticed [Figure 2.5]

In case 6 with $\lambda = 0.69$, the computation converges after 12 iterations and significantly better orthogonality is noticed. [Figure 2.8]

In case 8 with $\lambda = 50$, the computation converges after 17 iterations, less orthogonal grids noticed [Figure 2.11]

Larger $\lambda$ such as $\lambda = 600$ and $\lambda = 900$ were run but the orthogonality result is no better than when $\lambda = 50$, so they are not included in this thesis.

Computation converges after 17 iterations for higher $\lambda$.

The smaller $\lambda$ is and the larger $\mu$ is the more orthogonal the grid. $\mu$ is between 0-1, if not distortion occurs.

The boundary type as evident from results is the Neumann boundary condition. This gave good results.

With this method we don’t have to worry about grid folding as some techniques had because we have control of the cell size through the monitor function, $f$. We enforce the Jacobian determinant of the grid generated to be strictly positive through the control of a monitor function $f$. Unlike what has been done before we don’t have to add the volume functional, $I_v$ to generate good results.

3. Image Registration

The image registration problem consists of three major components:

1. Transformation models: Rigid, Non-rigid
2. Similarity measures: Intensity-based, Geometry-based
3. Optimization methods: Gradient Descent, Levenberg-Marquardt Optimization, Downhill Simplex method, Deterministic Annealing.

3.1. Transformation Models

Transformation models serve for two purposes. They control how image features can be moved relative to one another to improve the image similarity, secondly they interpolate between those features.

3.1.1. Rigid Image Registration

Rigid registration (Affine) is composed solely of a global rotation, translation, scaling and projection.

Translation: moving image from one position to another position
Rotation: changing the angle of image
Scaling: increasing the size of the actual image
Projection: representing the image on a plane as it would look from a particular direction

In Rigid Image registration angles and distances between points are preserved. Rigid transformations are global and linear; hence they can be represented by matrices.

As an example let us consider a 2D scaling from the origin.
Let point \( P \) be defined as \( (x, y) \).
Perform a scale (stretch) to point \( P' (x', y') \) by a factor \( s_x \) along the x-axis, and \( s_y \) along the y-axis.
\[
x' = s_x x, \quad y' = s_y y
\]
Define the matrix,
\[
S = \begin{bmatrix}
s_x & 0 \\
0 & s_y
\end{bmatrix}
\]
Now,
\[
P' = SP \quad \text{or} \quad \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} s_x & 0 \\ 0 & s_y \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}
\]

3.1.2. Non-Rigid Image Registration
Non-rigid image registration refers to a class of methods where the images to be registered have non-linear geometric differences.
A non-rigid transformation does not preserve the straightness of lines and in general, maps a line into a curve.
In this paper, our problem concerns only the Non-Rigid Image registration. Some non-rigid image registration techniques are viscous fluid algorithm, optical flow methods, thin-plate spline and cubic B-spline as discussed in [17].

3.2. Similarity Measures
Similarity metric is divided into two major components as intensity-based and landmark (geometry) based similarity metric.

3.2.1. Intensity-based Registration
A straightforward approach is based on the minimization of the so-called sum of squared differences (SSD); cf., e.g., [13] or [34].

Definition 5.2.1 Let \( d \in \mathbb{N} \) and \( R, T \in \text{Im} g(d) \).
The sum of squared differences (SSD) distance measure \( D_{\text{SSD}} \) is defined by
\[
D_{\text{SSD}}: \text{Im} g(d)^2 \rightarrow \mathbb{R},
\]
\[
D_{\text{SSD}}[R, T] := \frac{1}{2} \| T - R \|_{L_2}^2 = \frac{1}{2} \int_{\mathbb{R}^d} (T(x) - R(x))^2 \, dx
\]
For a transformation \( \Phi : \mathbb{R}^d \rightarrow \mathbb{R}^d \) we also define:
\[
D_{\text{SSD}}[R; T; \Phi] = D_{\text{SSD}}[R, T \circ \Phi] = \frac{1}{2} \int_{\mathbb{R}^d} (T(\Phi(x)) - R(x))^2 \, dx
\]
In this dissertation we will concentrate on minimizing SSD.

3.2.1.1. Mutual Information-based Registration (MI)
Since 1995, mutual information has been used in image registration. This measure has its roots in information theory and has demonstrated its power and robustness for use in multi-modality registration.
It was proposed independently by Viola in [35] and Collignon et al in [36] and has been used since by Kim et al in [37] and many others.
The basic idea is the maximization of the so-called mutual information of the images with respect to the transformation. Mutual information is an entropy-based measure; it measures a statistical dependence between the intensity of corresponding voxels as opposed to a functional dependency. This method and others are discussed further in [17].

3.3. Optimization Methods
Finding the minimum of dissimilarity measure or the maximum of similarity measure is a multidimensional optimization problem, where the number of dimensions corresponds to the degrees of freedom of the expected geometrical transformation.
The optimization problem for nonlinear registration is ill-posed for reasons which will be discussed in future chapters; we add regularization terms or penalty terms next to the dissimilarity measure term to be minimized, which interconnects the transformation and data to be transformed [38].
These two terms together form the cost function (energy) associated with the registration and the aim of the optimization methods is to minimize it.

Let us define the registration problem as:
Find a transformation \( \Phi(x) = u(x) + x \), such that
\[
J[u] = D[R, T; u] + \alpha S[u] \quad \text{is minimized.}
\]
Here \( D \) is a distance measure; \( S \) is a regularizing term (smoother) for the displacement \( u \) and \( \Phi(x) \) denotes the non-rigid transformation which equates to a translation of every pixel \( x \) in the template image by a certain displacement defined by the displacement field \( u(x) \).
We use parameters to control the strength of the smoothness of the displacement versus the similarity of the images. Regularization terms are added often to handle
folding, cracks or other unwanted deformations due to arbitrary transformations.

Typical regularizers are fluid, elastic, diffusive and curvature smoother. The shortcomings of adding these terms are discussed in [17].

4. Optimal Control Approach of an Image Registration Problem

In this chapter we will discuss the optimal control approach to non-rigid image registration with the intensity based method {Sum of Squared Differences (SSD)} using the grid deformation method and then define the optimal control problem and then discuss about proving the existence of optimal solutions, existence of Lagrange multipliers without the use of an ODE constraint, which was used in [39].

An ODE constraint is used in [39] to find $\Phi$, the time-dependent mapping but Chih-yao hsieh in [40] showed that we don’t need that extra constraint.

In [39], an existence of solution of the ODE constraint is proved using the Peano existence theorem and two well-known results in functional analysis. This helps in proving that an optimal solution exists.

4.1. Grid Deformation Method

In order to find an optimal transformation that minimizes the dissimilarity between the transformed image and the Reference image, we adopt the grid deformation method [41], which is essential in constructing differentiable and invertible transformations to solve mesh adaptation problems.

The deformation method has its origin in differential geometry [42]. It was reformulated for grid generation as discussed in chapter 1.

The grid deformation method gives direct control over the cell size of the adaptive grid and determines the node velocities directly. A great advantage of this method is that it avoids grid folding by enforcing the Jacobian determinant of the grid generated to be strictly positive through the control of a monitor function $f$. So, unwanted registration results are avoided.

4.1.1. Formulation

A transformation is defined in a two-step manner. Firstly, a given function is used to construct a vector field that satisfies a div-curl system and secondly, this vector field is used to generate a transformation that moves the grid.

Now, let us simplify the grid deformation method by letting $h(t,x) = h(1,x) = u(x)$ ([39]), for our image registration problem.

4.2. Set-up of the Cost (Objective) Functional

Let $\Omega \subset \mathbb{R}^2$ be a bounded domain that is a convex polygon [to simplify analysis and the numerical scheme] or that has a $C^{1,1}$ boundary ($\partial \Omega$).

Given:
- $R(x)$: Reference image fixed
- $S(x)$: {differentiable and invertible} is a transformation from $\Omega \to \Omega$

We assume that $S \approx \text{Id}$.

Define: $T(x) = R(S(x))$ i.e. we use $S$ to define the template image.

We register $T(x)$ to $R(x)$ by $\Phi(x)$, so that $R(x) \approx T(\Phi(x))$

$$\Phi(x) = u(x) + x$$

$$R(x) \approx T(\Phi(x)) = R(S(\Phi(x)))$$

$$\Rightarrow S(\Phi(x)) = x$$

$$\Rightarrow \Phi(x) = S^{-1}(x)$$

![Figure 4.1. Illustration of Registration goal](image)

Where,

$$\Phi(x) = u(x) + x = (u_1(x_1, x_2) + x_1, u_2(x_1, x_2) + x_2)$$

$x \in \Omega$

$$R(x), T(x) \in H^1(\Omega) \times H^1(\Omega)$$

$$u(x) \in H^2(\Omega) \times H^2(\Omega)$$

Let us make use of the Sobolev space defined as:

$$H^m(\Omega) = \{ p \in L^2(\Omega); D^\alpha p \in L^2(\Omega) \text{ for } 0 \leq |\alpha| \leq m \}$$

Let $H^m(\Omega) \times H^m(\Omega)$ denote the corresponding space of vector-valued functions each of whose components belong to $H^m(\Omega)$

For both scalar and vector-valued functions, let $\| \cdot \|_m$ denote the corresponding norm and inner
Optimal Control Approach to Image Registration

In order for the transformed template image to be close to the reference image, we seek a mapping \( \Phi(x) \) that minimizes the \( L^2 \) norm of the difference between \( T(\Phi(x)) \) and \( R(x) \) over \( \Omega \).

We define the functional:

\[
J(u,f,g) = \frac{1}{2} \|T(\Phi(x)) - R(x)\|^2 + \frac{1}{2} d_1 \|u\|^2 + \frac{1}{2} d_2 \|g\|^2 
\]

(4.2.1.1)

Where, \( d_1, d_2 \) are penalty parameters. Then we minimize

\[
\int_\Omega \left[ T(\Phi(x)) - R(x) \right]^2 dx
\]

Subject to:

\[
\text{div } u = \nabla \cdot u = f - 1 \quad \text{in } \Omega \\
\text{curl } u = \nabla \times u = g \quad \text{in } \Omega
\]

According to the Helmholtz Decomposition Theorem, a vector field (\( u \) in our case) can be constructed with both a specified divergence and curl as above.

The theorem below will be used in the proof for the existence of optimal solutions.

Theorem 4.2.1: Bounded Convergence Theorem

Let \( \{u^n\} \) be a bounded sequence of measurable functions defined on a set \( \Omega \) of finite measure and suppose that

\[
\lim_{n \to \infty} u^n(x) = f(x) \quad \forall x \in \Omega
\]

Then

\[
\lim_{n \to \infty} \int_\Omega f^n = \int_\Omega f
\]

Theorem 4.2.2: Linearization of \( u \) at identity

Let \( \Omega \subset \mathbb{R}^2 \) be a domain. For \( \Phi(x) = u(x) + x = (u_1(x_1, x_2) + x_1, u_2(x_1, x_2) + x_2) \)

The equality \( J(\Phi) \equiv 1 + \text{div}(\varepsilon u) + O(\varepsilon) \) holds for every \( \varepsilon > 0 \) in \( \Omega \).

Proof:

\[
J(\Phi) = \frac{\partial u_1}{\partial x_1} \frac{\partial u_1}{\partial x_2} = \left| \begin{array}{cc}
\varepsilon u_{1x_1} & \varepsilon u_{1x_2} \\
\varepsilon u_{2x_1} & 1 + \varepsilon u_{2x_2}
\end{array} \right|
\]

\[
= (1 + \varepsilon u_{1x_1}) (1 + \varepsilon u_{2x_2}) - (\varepsilon u_{1x_2})(\varepsilon u_{2x_1}) = \\
1 + \varepsilon u_{2x_2} + \varepsilon u_{1x_1} + \varepsilon^2 u_{1x_2} u_{2x_2} - \varepsilon^2 u_{1x_2} u_{2x_1} \\
1 + \varepsilon u_{2x_2} + \varepsilon u_{1x_1} + \varepsilon^2 (u_{1x_2} u_{2x_2} - u_{1x_2} u_{2x_1}) \equiv \\
1 + \varepsilon u_{2x_2} + \varepsilon u_{1x_1} + O(\varepsilon^2) = 1 + \text{div}(\varepsilon u) + O(\varepsilon^2)
\]

We then seek controls \( f \) and \( g \) and state \( \Phi \) such that \( \|T(\Phi(x)) - R(x)\|^2 \) in (6.2.1.2) is minimized subject to the constraints:

\[
\nabla \cdot u = f - 1 \quad \text{in } \Omega \\
\nabla \times u = g \quad \text{in } \Omega
\]

(4.2.1.2)

Since our objective is to minimize \( \int_\Omega \left[ T(\Phi(x)) - R(x) \right]^2 dx \) in (4.2.1.1); but doing this can result in unbounded optimal control functions due to the lack of any explicit dependence of the cost functional on the controls.

In order to counter this, we penalize the cost functional with \( H^1 \) - norms of the controls \( f \) and \( g \). With this, we can show the existence of optimal solutions.

According to [39] using a weaker \( L^2 \) - norm penalization, existence of optimal solutions has not yet been proven and computational studies indicate that \( L^2 \) - norm penalization may not be sufficient to guarantee the existence of optimal solutions.

Let us define the product Hilbert Spaces:

\[
V = [H^2(\Omega)]^2 \times H^1(\Omega) \\
W = H^1(\Omega) \times H^1(\Omega) \times \mathbb{R} \times H^2(\Omega) \times L^2(\Omega)
\]

And define the constraint operator:

\[
M : V \to W \text{ as follows:}
\]

\[
\{u, f, g\} = 0 \text{ is the constraint equation} \\
\]

(4.2.1.3)

We recognize (4.2.1.3) as a weak formulation of the constraint equations (4.2.1.2).

The admissible set of all solutions is defined by:

\[
U_{ad} = \{(u,f,g) \in V : J(u,f,g) \text{ is bounded and } M(u,f,g) = 0\} 
\]

(4.2.1.4)

Then the optimal control problem is given by:

Find \( (u,f,g) \in U_{ad} \) which minimizes

\[
\int_\Omega \left[ T(\Phi(x)) - R(x) \right]^2 dx 
\]

(4.2.1.5)

4.3. Existence of Optimal Solutions

We next show the existence of solutions of the optimal control problem (4.2.1.5). Under the assumptions we have
made about $\Omega \subset \mathbb{R}^2$, it is well known that given $f,g \in H^1(\Omega)$, \( \exists \ a \ u \in [H^2(\Omega)]^2 \) satisfying:
\[
\begin{align*}
\nabla \cdot u &= f - 1 \quad \text{and} \quad \nabla \times u &= g \quad \text{in} \ \Omega, \\
n \cdot u &= 0 \quad \text{on} \ \partial \Omega
\end{align*}
\]
Since $u \in [H^2(\Omega)]^2$, by the Sobolev imbedding theory, $u$ is bounded and Hölder-continuous i.e. \( \forall x, y \in \Omega \) for some \( K_1 > 0 \), \( |u(x)| \leq K_1 \) and for some constant \( K \):
\[
|u(x) - u(y)| \leq K_2 |x - y|^2, \quad \text{for} \ 0 < \lambda < 1
\]
Also, the div-curl system described in (4.2.1.1) is elliptic in the sense of Petrovski, with this we can also obtain a uniform bound for $u$.

**Lemma 4.3.1: Petrovski Ellipticity**

Consider the div-curl system in 2D,
\[
\begin{align*}
\nabla \cdot u &= f - 1 \\
\nabla \times u &= g \quad \text{in} \ \Omega
\end{align*}
\]
with \( n \cdot u = 0 \) \( \text{in} \ \partial \Omega \)

Let \( f^* = f - 1 \)

M is elliptic in the sense of Petrovski (i.e. regular Ellipticity)
\[
N(u) = \begin{pmatrix} u_{x_1} + u_{x_2} \\ u_{y_1} - u_{x_2} \end{pmatrix} = \begin{pmatrix} f^* \\ g \end{pmatrix}
\]

Where,

State variable: \( u(x) = (u_1(x_1, x_2), u_2(x_1, x_2)) \) is a vector valued function in \([H^1(\Omega)]^2\)

Control variables: \( f^* \) and \( g \) \{scalar valued functions\}

in \([H^1(\Omega)]^2\)

We can write the above as a linear system:
\[
A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}
\]
n is a unit outward normal to \( \Omega \)

R is a full rank \( 2 \times 2 \) matrix

The boundary condition satisfies the complementing conditions of Agmon, Douglis and Nirenberg in [45].

\[
N(u) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} u_{1x_1} \\ u_{2x_1} \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} u_{1x_2} \\ u_{2x_2} \end{pmatrix} = \begin{pmatrix} f^* \\ g \end{pmatrix}
\]

Now,

Ellipticity in the sense of Petrovski requires that:
\[
\det(A\lambda + B\eta) = 0 \quad \iff \quad \lambda = 0 \quad \text{and} \quad \eta = 0
\]
for \( (\lambda, \eta) \in \mathbb{R}^2 \)

Now,

Assume \( \lambda \neq 0 \) and \( \eta \neq 0 \)

\[
\det(A\lambda + B\eta) = \det\left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right) = \lambda^2 + \eta^2 \neq 0
\]

As a consequence of the div-curl system being regular elliptic, we have that there exists a constant \( C > 0 \) such that:
\[
\frac{1}{C} \|v\|_0 \leq \|\nabla \cdot v\|_0 + \|\nabla \times v\|_0 \leq C\|v\|_0
\]
with \( v \in H^1(\Omega) \)

If \( f \) satisfies \( \int_\Omega (f - 1)\nabla \cdot v = 0 \) then \( \exists \) a unique solution \( u \in [H^2(\Omega)]^2 \) that satisfies the above approximation.

As a consequence, we deduce a uniform bound for $u$.

**Theorem 4.3:** There exists a solution \( (u, f, g) \in U_{ad} \) for the optimal control problem (4.2.1.5)

**Proof**

We first choose $f=1$ and $g=0$. Then, we immediately see that in constraint $\nabla \cdot u = 1 - 1 = 0$

\[ \Rightarrow u \quad \text{is a constant i.e.} \quad u = 0 \]

\[ \Rightarrow U_{ad} \neq \emptyset \quad \text{because} \quad (0,1,0) \in U_{ad} \]

Let \( \{u^n, f^n, g^n\} \subset U_{ad} \) denote a minimizing sequence i.e. we have that
\[
\lim_{n \to \infty} J(u^n, f^n, g^n) = \inf_{(u, f, g) \in U_{ad}} J(u, f, g)
\]

By (4.2.1.4), \( f^n \) and \( g^n \) are bounded in \( H^1(\Omega) \), \( \|f^n\|, \|g^n\| \) are bounded.

Since \( u^n \in \left[ H^2(\Omega) \right]^3 \), \( u^n \) is bounded i.e. \( \|u^n\| \leq K \) for some \( K > 0 \).

Then we can extract subsequences such that:

\[
\begin{align*}
  u^n &\to \tilde{u} \quad \text{weakly in} \quad \left[ H^2(\Omega) \right]^3 \\
  f^n &\to \tilde{f} \quad \text{weakly in} \quad H^1(\Omega) \\
  g^n &\to \tilde{g} \quad \text{weakly in} \quad H^1(\Omega)
\end{align*}
\]

for some \( \left( \tilde{u}, \tilde{f}, \tilde{g} \right) \in V \)

Next step is to show that the limit \( \left( \tilde{u}, \tilde{f}, \tilde{g} \right) \in V \) satisfies the constraint equations in (4.2.1.3)

\[
\nabla \cdot u - f + 1, \xi = 0 \quad \forall \xi \in H^{-1}(\Omega)
\]

PROOF

For a given \( f \in H^1(\Omega) \), pick \( \xi \in H^{-1}(\Omega) \)

Let us show that:

\[
\langle \nabla \cdot \tilde{u}, \xi \rangle = \langle \tilde{f} - 1, \xi \rangle
\]

By (4.2.1.6),

\[
\begin{align*}
\int_\Omega (\nabla \cdot u^n - f^n + 1 - (\nabla \cdot \tilde{u} - \tilde{f} + 1)) \cdot \xi \, d\Omega &= \\
\int_\Omega (\nabla \cdot u^n - \nabla \cdot \tilde{u} - f^n + \tilde{f}) \cdot \xi \, d\Omega &= \\
\int_\Omega (\nabla \cdot u^n - \nabla \cdot \tilde{u} - (f^n - \tilde{f})) \cdot \xi \, d\Omega &\to 0
\end{align*}
\]

\( \tilde{u} \) and \( \tilde{g} \) satisfy \( \langle \nabla \times u - g, \eta \rangle = 0 \quad \forall \eta \in H^{-1}(\Omega) \)

PROOF

\[
\langle \nabla \times \tilde{u}, \eta \rangle = \langle \tilde{g}, \eta \rangle
\]

By (4.2.1.6),

\[
\begin{align*}
\int_\Omega (\nabla \times u^n - g^n - (\nabla \times \tilde{u} - \tilde{g})) \cdot \eta \, d\Omega &= \\
\int_\Omega ((\nabla \times u^n - \nabla \times \tilde{u}) - (g^n - \tilde{g})) \cdot \eta \, d\Omega &\to 0
\end{align*}
\]

Now, let us show that \( \tilde{f} \) satisfies

\[
\sigma \int_\Omega (f - 1) \, d\mathbf{x} = 0, \quad \forall \sigma \in \mathcal{R}
\]

PROOF

If \( \sigma = 0 \), trivially \( \tilde{f} \) satisfies the constraint. Assume \( \sigma \neq 0 \)

Since \( f^n \) is bounded, therefore \( f^n - 1 \) is bounded

\[
0 = \lim_{n \to \infty} \int_\Omega (f^n - 1) = \int_\Omega \lim_{n \to \infty} (f^n - 1) = \int_\Omega (\tilde{f} - 1)
\]

by the bounded convergence theorem

[see Theorem 4.2.1] and (4.2.1.6)

Since \( \Omega \) is a convex polygon in \( \mathbb{R}^2 \), its interior is a convex set, which is closed and bounded.

This implies that \( J(\cdot) \) is a convex functional.

Since in general it is difficult to prove weak lower semi-continuity, convexity is a sufficient condition for this.

This tells us that \( J(\cdot) \) is weakly lower semi-continuous.

We know \( n \cdot u = 0 \), since \( u^n \) is bounded. By the choice of \( u^n \)'s we have \( n \cdot u^n = 0 \quad \forall n \)

We have \( \lim_{n \to \infty} n \cdot u^n = n \cdot \tilde{u} = \lim_{n \to \infty} 0 \) from (4.2.1.6).

Thus, \( \left( \tilde{u}, \tilde{f}, \tilde{g} \right) \subseteq V \) satisfies \( M(u, f, g) = 0 \) i.e.

\[
\lim_{n \to \infty} J(u^n, f^n, g^n) = J(\tilde{u}, \tilde{f}, \tilde{g}) = \inf_{(u, f, g) \in U_{ad}} J(u, f, g)
\]

Thus, we have shown that an optimal solution belonging to \( U_{ad} \) exists.

4.4. Non-Uniqueness of Solutions

Since the image registration problem is ill-posed, solutions are not unique. This was discussed in Chapter 1; also see Figure 1.2 for illustration.

4.5. Existence of Lagrange Multipliers

Now we want to use the Lagrange multipliers rule to turn the constrained minimization problem in (4.2.1.1) into an unconstrained problem and then to derive an optimality system. So, we show the existence of proper Lagrange multipliers for any \( (u, f, g) \).

Let \( M \) be the constraint equations from \( V \) to \( W \) defined as in (4.2.1.3). Let \( M \left( u, f, g \right) \subseteq L(V; W) \) be the first Frechét derivative of \( M \) at \((u, f, g)\) defined as \( M(u, f, g) = (\tilde{u}, \tilde{f}, \tilde{g}) \) where \( \tilde{u}, \tilde{f}, \tilde{g} \) are defined as in (4.2.1.3).

Now, let us show that \( \tilde{f} \) satisfies

\[
\sigma \int_\Omega (f - 1) \, d\mathbf{x} = 0, \quad \forall \sigma \in \mathcal{R}
\]

PROOF
Next we prove that a suitable Lagrange multiplier exists. To prove this we need to use the following theorem concerning the existence of Lagrange multipliers on Banach spaces [43].

**Theorem 4.5.1:** Let $V_1$ and $V_2$ be two Hilbert spaces, $F$ a functional on $V_1$, and $G$ a mapping from $V_1$ to $V_2$. Assume $\tilde{u}$ is a solution of the following constrained minimization problem:

Find $u \in V_1$ that minimizes $F(u)$ subject to $G(u) = 0$. Assume further that the following conditions are satisfied:

1. $\nabla \cdot F(\tilde{u})$ is Frechet-differentiable at $\tilde{u}$;
2. $G(\tilde{u}): V_1 \rightarrow V_2$ is onto;
3. $G'(\tilde{u}) : V_1 \rightarrow V_2$ is onto

Then, there exists a $\mu \in (V_2')^*$ such that

$$F'(\tilde{u})v - \langle \mu, G'(\tilde{u})v \rangle = 0, \quad \forall v \in V_1$$

**Proof:**

The proof can be found in [43], Theorem 43.19 and [44].

We have found a $(\tilde{u}, \tilde{f}, \tilde{g})$ satisfying (4.2.1.7)

Now, we will prove that a non-zero Lagrange multiplier exists by following Theorem 4.5.1. We construct the Lagrangian functional as follows:

$$L(u, f, g, \Theta_1, \Theta_2) = J(u, f, g) + \int_\Omega \Theta_1 (\text{div} u - (f - 1)) + \int_\Omega \Theta_2 (\text{curl} u - g)$$

And then taking the derivative we have:

$$\dot{L}(u)v = J(u)v + (M'(u)v) \cdot (v_1, v_2)$$

Where,

$$M'(u)v = (\text{div} u - (f - 1), \text{curl} u - g) \cdot (v_1, v_2)$$

and $\Theta_1, \Theta_2$ are Lagrange multipliers.

The above is used in Theorem 4.5.2 below to prove the existence of multipliers, still following Theorem 4.5.1.

**Theorem 4.5.2:** Let $(\tilde{u}, f, g) \in V$ denote an optimal solution to the optimal control problem previously discussed. Then there exists a non-zero Lagrange multiplier $\Theta \in W^*$ such that:
\[ J'(u, f, g; \tilde{u}, \tilde{f}, \tilde{g}) + \langle M'(u, f, g; \tilde{u}, \tilde{f}, \tilde{g}) \Theta, 0 \rangle = 0 \quad \forall (\tilde{u}, \tilde{f}, \tilde{g}) \in V \]

which contradicts the hypothesis that \((u, f, g)\) is an optimal solution to (4.2.1.5).

To conclude, we use the Hahn-Banach Theorem. Since, \( \text{Im}(N'(u, f, g)) \) is closed in \( \mathcal{R} \times W \), for any \((a, F) \in \text{Im}(N'(u, f, g))\) \( \exists \) a non-zero \( \Theta \in (\mathcal{R} \times W)^* \) satisfying \( \langle (a, F), (\tilde{a}, \Theta) \rangle = 0 \)

Suppose \( \tilde{a} = 0 \), then \( \langle F, \Theta \rangle = 0 \quad \forall F \).

Therefore, \( \tilde{a} \neq 0 \). Without loss of generality we can set \( \tilde{a} = \mathbf{1} \) and therefore the theorem holds.

So, the penalized Lagrangian functional is:

\[ L(u, f, g, \xi, \eta, \nu, \sigma, v) = J(u, f, g) + \int_{\Omega} (\nabla \cdot u - f + 1)\xi \, dx + \int_{\Omega} (\nabla \times u - g)\eta \, dx + \sigma \int_{\Omega} (f - 1) \, dx + \int_{\Omega} (n \cdot u) \, vds \]

\[ W^* = H^{-1}(\Omega) \times H^{-1}(\Omega) \times \mathcal{R} \times H^{-3}(\partial \Omega) \]

### 4.6. Optimality System

Solution of the Lagrangian functional \( L \) above is called the optimality system which consists of state equations, costate equations and the optimality conditions.

First we include a Lemma that we will use.

Lemma 4.5.2: The equalities hold for any scalar function \( h \) defined on \( \Omega \) and a vector \( \nu \).

\[ \int_{\Omega} \nabla \cdot (h \nu) = \int_{\Omega} \nabla h \cdot \nu + \int_{\Omega} h \nabla \cdot \nu \]

\[ \int_{\Omega} \nabla \cdot (h \nu) = \int_{\Omega} (h \nu) \cdot \bar{n} = 0, \text{ if } h = 0 \text{ on } \partial \Omega \]

Where \( \bar{n} \) is the outward and unit normal vector.

State Equations: The state equations are obtained by solving the equations \( L_u = 0 \) and \( L_{g} = 0 \)

\[ L_u = \frac{d}{d \xi} |_{\xi = 0} L[\xi + \varepsilon \delta \xi] = \frac{d}{d \xi} \int_{\Omega} (\nabla \cdot u - f + 1)(\xi + \varepsilon \delta \xi) = \]

\[ = \int_{\Omega} (\nabla \cdot u - f + 1)\delta \xi = 0 \quad \forall \delta \xi \]

Then,

\[ \text{div}u = \nabla \cdot u = f - 1 \quad (4.2.1.8) \]

\[ L_g = \frac{d}{d \eta} |_{\eta = 0} L[\eta + \varepsilon \delta \eta] = \frac{d}{d \eta} \int_{\Omega} (\nabla \cdot u - g)(\eta + \varepsilon \delta \eta) = \]

\[ = \int_{\Omega} (\nabla \cdot u - g)\delta \eta = 0 \quad \forall \delta \eta \]

Then,
\[
\text{curl } u = \nabla \times u = g \quad (4.2.1.9)
\]

The state equations are given by:
\[
div u = \nabla \cdot u = f - 1
\]
\[
\text{curl } u = \nabla \times u = g
\]

Costate Equations: The costate equations are obtained by solving the equations \( L_{u_1} = 0 \) and \( L_{u_2} = 0 \)

\[
L_{u_i} = \frac{d}{dx} \left. \left[ \frac{1}{2} \int_{\Omega} T(x + (u_i + \varepsilon \delta u_i, u_2)) - R(x) \right]^2 \right|_{x=0}
+ \int_{\Omega} \xi \left( \text{div} (u_i + \varepsilon \delta u_i, u_2) \right) - f + 1) dx
+ \int_{\Omega} \eta \left( \text{curl} (u_i + \varepsilon \delta u_i, u_2) \right) - g dx
\]

\[
= \int_{\Omega} T(x + u(x) - R(x)) \nabla \delta u_1 + \int_{\Omega} \xi \delta u_1 + \int_{\Omega} \eta \left( \text{curl} (u_i, u_2) \right) - g dx
\]

This gives us the first costate equation
\[
\nabla \cdot (\xi, -\eta) = (T - R) \nabla \delta u_1 \quad (4.2.2.0)
\]

Optimality conditions:
\[
d_1 f = -\xi - \sigma
\]
\[
d_2 g = \eta \quad (4.2.2.3)
\]

Now we write the optimality system as follows:

State Equations:
\[
div u = \nabla \cdot u = f - 1
\]
\[
\text{curl } u = \nabla \times u = g
\]

Costate Equations:
\[
\nabla \cdot (\xi, -\eta) = (T - R) \nabla \delta u_2
\]
\[
\nabla \cdot (\eta, \xi) = (T - R) \nabla \delta u_2
\]

Optimality conditions:
\[
d_1 f = -\xi - \sigma
\]
\[
d_2 g = \eta
\]

Details of how to solve these numerically by multi-grid optimization were discussed in [17]. It is shown how we can solve these equations.

5. Conclusions

In this paper we have shown that orthogonal grids can be obtained with a regularization term smoothness functional which has a unique solution. We do this without the volume or area functional. The smoothness functional is added because the orthogonality problem is ill-posed.

The only term we add is a penalty term \{smoothness functional\} that regularizes the functional.

We optimize the orthogonality of grid by minimizing an orthogonal functional, this procedure involves solving for a displacement \( u \) which minimizes the functional \( I \) and satisfies the Euler-Lagrange equations.
The concern of grid folding is prevented by the grid deformation method as discussed previously and good results were obtained for small choices of $\lambda$ and large $\mu$. The orthogonal grids should give more accurate solutions to Partial Differential equations.

REFERENCES

[1] G. Liao and D. Anderson, “A New Approach to grid Generation”, Appl. Anal. 44, (1992).
[2] G. Liao and J. Su, “A direct method in Dacorogna-Moser’s approach of grid generation problems”, Appl. Anal., 45, (1993).
[3] G. Liao and J. Su, “Grid Generation via Deformation”, Appl. Math. Lett., 5, (1992).
[4] G. Liao, T. Pan, and J. Su, “A Numerical Grid Generator based on Moser’s Deformation method”, Numer. Meth. PDE, 10, (1994).
[5] F. Liu, S. Ji, and G. Liao,” An Adaptive Grid Method with cell-volume control and its applications to Euler Flow calculations”, SIAM J. Sci. COMPUT., 20, (1998).
[6] G. Liao and J. Su, “A moving grid method for (1+1) dimension”, Appl. Math Lett., 8, (1995).
[7] B. Semper and G. Liao,” A moving grid finite element method using grid deformation”, Numer. Meth. PDE, pp. 11, (1995).
[8] G. Dela Pena “Adaptive grid generation”, PhD thesis, UT, Arlington, (1998).
[9] B. Dacorogna and J Moser,” On a PDE involving the Jacobian Determinant”, Ann. Inst H Poincare, 7, (1990).
[10] G. Liao, Gary dela Pena, Guojao Liao,” A Deformation method for moving grid generation”.
[11] M. Akinlar “A new method for Non Rigid Registration of 3D images”, PhD thesis, UT, Arlington, (2009).
[12] J. Modersitzki, Numerical Methods for Image Registration, Numerical Mathematics and Scientific Computation, Oxford Science Publications, pp. 14-15, (2004).
[13] [13] L. G. Brown. A survey of image registration techniques. ACM Computing Surveys 24(4), pp. 325-376, (1992).
[14] J. B. A. Maintz and M.A. Viergever, A survey of medical image registration, Medical image Analysis 2(1), pp. 1-36, (1998).
[15] C.R. Maurer, and J. M. Fitzpatrick, A review of medical image registration. In Interactive Image-Guided Neurosurgery, pp. 17-44. American Association of Neurological Surgeons, Park Ridge, (1993).
[16] P.A. Van den Elsen, E.-J.D. Pol, and M.A., Viergever, Medical image matching- a review with classification. IEEE Engineering in Medicine and Biology, 26-38, (1993).
[17] M. Akinlar “A new method for Non Rigid Registration of 3D images”, PhD thesis, UT, Arlington, (2009).
[18] Joe F. Thompson, Z. U. A. Warsi, C. Wayne Mastin, Numerical Grid Generation, Foundations and Applications, Elsevier Science Publishing Co., (1985).
[19] P. Das, A higher order difference method for singularly perturbed parabolic partial differential equations, Journal of Difference Equations and Applications, 24, 3, 452-477, doi: 10.1080/10236198.2017.1420792, 2018.
[20] P. Das and J. Vigo-Aguiar, Parameter uniform optimal order numerical approximation of a class of singularly perturbed system of reaction diffusion problems involving a small perturbation parameter, Journal of Computational and Applied Mathematics, doi: 10.1016/j.cam.2017.11.026, 2017.
[21] P. Das and S. Natesan, Numerical solution of a system of singularly perturbed convection diffusion boundary value problems using mesh equidistribution technique, Australian Journal of Mathematical Analysis and Applications, 10(1), 1-17, 2013.
[22] M. Chandru, T. Prabha, P. Das, V. Shanthi, A numerical method for solving boundary and interior layers dominated parabolic problems with discontinuous convection coefficient and source terms, Differential Equations and Dynamical Systems, doi: 10.1007/s12591-017-0385-3, 2017.
[23] P. Das, An a posteriori based convergence analysis for a nonlinear singularly perturbed system of delay differential equations on an adaptive mesh, Numerical Algorithms, doi: 10.1007/s11075-018-0557-4, 2018
[24] M. Chandru, P. Das and H. Ramos, Numerical treatment of two-parameter singularly perturbed parabolic convection diffusion problems with non-smooth data, Mathematical Methods in the Applied Sciences, doi: 10.1002/mma.5067, 2018.
[25] G. Liao, Variational Approach to Grid Generation, Numerical methods for Partial differential equations, 8, pp. 143-147., (1992).
[26] J. U. Brackbill and J. S. Saltzman,” Adaptive zoning for singular problems in two dimensions,” J. Comput. Phys. 46, pp. 342-368 (1982).
[27] S. Steinberg and P. Roache,” Variational grid generation,” Numer. Meth. Partial Diff. Eq. 2, pp 1-17 (1995).
[28] J. Castillo, S. Steinberg, and P. J. Roache, “Parameter estimation in variational grid generation,” Appl. Math. Comput. 28, pp. 1-23 (1988).
[29] R. Courant, Trans. Amer. Math. Soc. 50 (1941), pp. 40.
[30] J. F. Thompson, F. C. Thames, and C. W. Masten, J. Comput. Phys. 15 (1974), pp. 299
[31] G. Liao and H. Liu, “Existence and $C^{0,\alpha}$ regularity of minima of a functional related to the grid-generation problem”, Numerical methods for partial differential equations, 9, pp. 261-264, (1993).
[32] P. Das and S. Natesan, A uniformly convergent hybrid scheme for singularly perturbed system of reaction-diffusion Robin type boundary value problems, Journal of Applied
[33] P. Das and S. Natesan, Higher order parameter uniform convergent schemes for Robin type reaction diffusion problems using adaptively generated grid, International Journal of Computational Methods, 9(4), 2012, doi: 10.1142/S0219876212500521.

[34] M. Capek, Optimisation strategies applied to global similarity based image registration methods. In WSCG '99: the 7th International Conference in Central Europe on Computer Graphic, pp. 369-374, (1999).

[35] P.A. Viola, Alignment by Maximization of Mutual Information, PhD thesis, Massachusetts Institute of Technology. (1995).

[36] A. Collignon, A. Vandermeulen, P. Suetens, and G. Marchal, 3d multi-modality medical image registration based on information theory. Computational Imaging and vision 3, pp. 263-274, Kluwer Academic, Dordrecht. (1995).

[37] B. Kim, J. L. Boes, K. A., Frey and C.R. Meyer, Mutual information for automated unwarping of rat brain autoradiographs. Neuroimage 5, 31-40. Article No. N1960251. (1997).

[38] J. V. Hajnal, D. L. G. Hill, D. J. Hawkes, Medical Image Registration, CRC Press, Baton Rouge, Florida, ISBN 0-8493-0064-9. (2001).

[39] Max Gunzburger, Eunjung Lee, An Optimal Control Formulation of an Image Registration Problem, Journal of mathematical imaging and vision. (2008).

[40] C.-Y.Hsieh, Non-rigid image registration by the Deformation based grid generation, Ph.D thesis, UT, Arlington, (2008).

[41] P. Bochev, G. Liao, G. dela Pena, Analysis and computation of adaptive moving grids by deformation. Numer. Methods Partial Differential Equations 12(4), pp. 489-506 (1996).

[42] J. Moser, “Volume Elements of a Riemann Manifold”, Trans AMS, pp. 120, (1965).

[43] E. Zeidler, Nonlinear analysis and its applications, 62, pp. Springer, New York, (1998).

[44] E. Zeidler, Applied Functional Analysis ‘Main principles and their applications’, volume 109, 1995 by springer-Verlag NY inc

[45] Agmon, Douglis and Nirenberg, Communications on Pure and Applied Mathematics, Vol. XVII, pp. 35-92 (1964).