OPTIMAL CONTROL APPROACH
TO IMAGE REGISTRATION

by

STEPHEN TAIWO SALAKO

Presented to the Faculty of the Graduate School of

The University of Texas at Arlington in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT ARLINGTON

DECEMBER 2009

ACKNOWLEDGEMENTS

I thank God for allowing me to achieve this feat, which seemed so daunting at times and for giving me the gift of life to see this through.

I want to thank Dr. Guojun Liao for being my supervising advisor and for his encouragement, patience and guidance during my doctoral study. I wish to thank Dr. Tuncay Aktosun, Dr. David Liu, Dr. Stephen Pankavich and Dr. Hristo Koujouharov for taking their precious time out to serve on my committee. I want to thank The University of Texas at Arlington Mathematics Department for helping me achieve this vision by giving me the opportunity to attend this great University and for assisting me financially through teaching Assistantships and doctoral scholarships during my graduate study. I would like to thank Dr. Li for his expertise in programming and all my UTA friends.

I want to thank God for my parents Professor Nathanael Salako and Mrs. Mobolanle Salako for their financial support, emotional support and encouragement throughout. You truly are part of the reason why this vision became reality and I am eternally grateful. I would also like to thank my sister, Nathalie for her continuous encouragement and moral boosting perception of me. If there was any doubt in me, ability wise, talking to her erases it all. I would like to thank my brothers David and Nathan. Grandpa, your prophecy has been fulfilled. I would like to thank my wife, Busola for her patience during the latter stages of my program. You truly are a gift from God and I love you with all my heart, thank you for seeing this through with me and thank you for bringing good fortune on my side.

November 17, 2009

ABSTRACT

OPTIMAL CONTROL APPROACH

TO IMAGE REGISTRATION

STEPHEN TAIWO SALAKO, PhD

The University of Texas at Arlington, 2009

Supervising Professor: Guojun Liao

The purpose of this dissertation is twofold: To present a new method of orthogonal grid generation and to investigate certain theoretical aspects of the optimal control approach to the image registration problem.

In the first part of this dissertation, we will present a variational method of orthogonal grid generation which is based on the deformation method and solving Euler-Lagrange equations. Although the concept of grid generation has been studied extensively, the generation of orthogonal grids is still one of the most challenging problems of the grid generation methods.

An orthogonal grid should offer significant advantages in the solution of systems of partial differential equations. Previous work requires a uniform grid to improve the orthogonality of grids. The grid deformation method provides size control via a monitor function but it has no control over gridline orthogonality.

Our approach is to improve orthogonality while providing size distribution by the deformation method.

In the second part, we replace the cost functional in the orthogonality problem with the sum of squared differences known as SSD to solve the Image registration problem.

Image registration is a significant part of image processing. It is a process of finding an optimal geometric transformation between corresponding pixels that minimizes the SSD.

In this optimal control approach, we use the grid deformation equations as constraints.

In this dissertation, we prove the existence of optimal solutions using the direct method in the calculus of variations; discuss the non-uniqueness of solutions and the existence of Lagrange multipliers of the optimal control problem for image registration using an abstract theorem concerning the existence of Lagrange multipliers on Banach spaces.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

Figure Page

LIST OF TABLES

CHAPTER 1

INTRODUCTION

In this chapter we describe the grid deformation method which forms the foundation for the optimization of grid orthogonality as well as the image registration problem. We also define what an image is and image registration.

1.1 Grid Deformation Method

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.

There are numerous approaches to grid generation as discussed in [1]. In this dissertation we use the deformation method to generate moving grids.

The grid deformation method is used for the construction of differentiable and invertible transformations to solve mesh adaptation problems. It can specify the Jacobian determinant of the grid transformation.

The idea of the deformation method is to move nodes with correct velocities so that the nodal mapping has a desirable Jacobian determinant. This gives precise volume control of the moving grids.

The static version of the grid deformation method was developed in [2], [3], [4]; it was improved in [5] and used with a finite-volume solver in flow calculations in [6]. A 2D version of the method was proposed in [7] and used with a discontinuous Galerkin finite-element method in solving a convection diffusion problem in [8].

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

Theorem 1.2: Let D be in $\Re^n$ be a bounded open set diffeomorphic to the unit ball B by a $C^{k+1,\alpha}$ diffeomorphism $\varphi : D \to B$ with $J(\varphi) > 0$ on closure of D and meas(D)= meas(B).Then there exists a $C^{k+1,\alpha}$ automorphism $F : \overline{D} \to \overline{D}$ satisfying $J(F(x)) = f(x)$, for x in D; $F(x) = x$ for x on ∂D.

Liao et al improved this theorem for the grid generation problem in [10]. There are three versions of the grid deformation method. We use one of the versions for the optimization of grid orthogonality. The versions are discussed in [11].

1.2 Image Registration

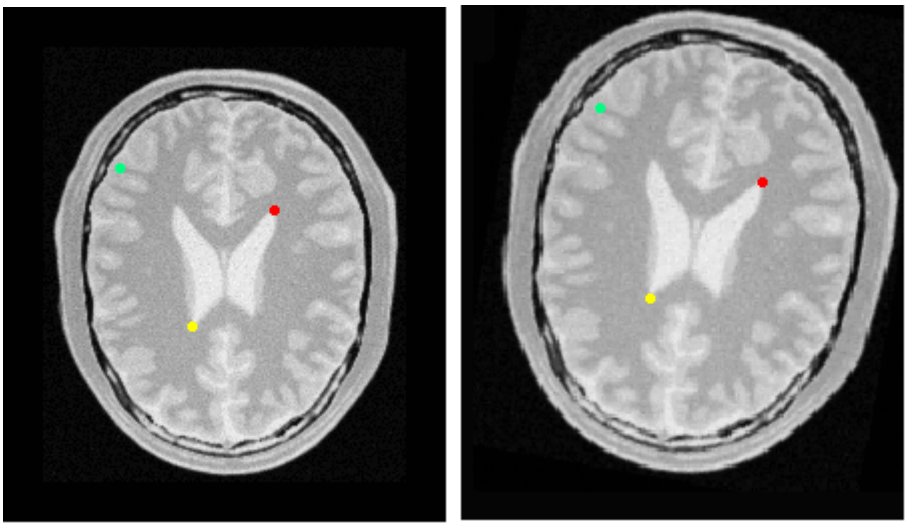
An image is a 2 dimensional array of pixels with assigned brightness values. A pixel is the smallest item of information in an image.

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

Some popular imaging modalities are x-ray, computed axial tomography (CAT), magnetic resonance imaging (MRI), positron emission tomography (PET), functional MRI (fMRI) and ultrasound (US).

Mathematically, 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 correspondence depends on the specific application.

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



(a) (b)

Figure 1.1 Template image (a) and Reference image (b)

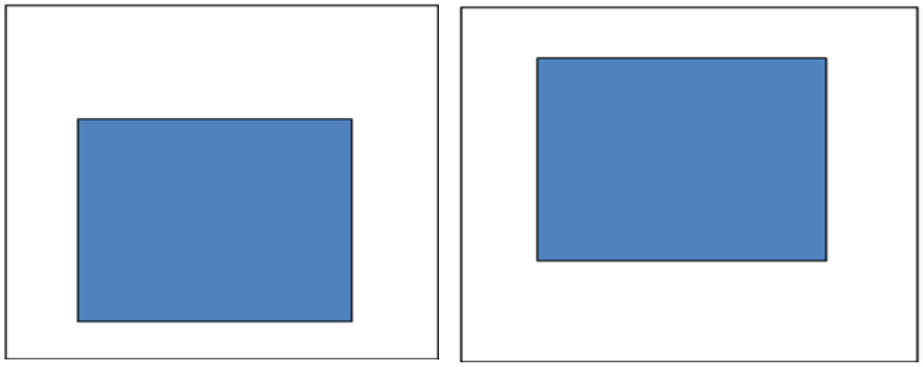
Notice that the reference image has undergone an affine deformation. Our goal is to warp the Template image to the Reference image by means of a transformation. The two images in Figure 1.1 are provided using the same imaging modality i.e. MRI.

Image registration is a problem often encountered in many applications areas like, for example, geophysics, computer vision and medical imaging. For an overview, see, [13], [14], [15] and [16].

Image registration is an ill-posed problem, which means that solutions are not unique and that small changes to the input images can lead to completely different registration results.

Non-uniqueness can be seen as follows:

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



(a) (b)

Figure 1.2 Template image (a) and Reference image (b)

For this pair, there are several solutions namely a pure translation, a rotation of 180° 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 shapes

(4.) Different sensors (multimodal analysis)

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 links between structural and functional images can be done.

CHAPTER 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 a penalty term. The goal is to improve orthogonality while keeping the size distribution of the grids.

We will discuss what has been done before and the significant improvement we have made in this dissertation. Formulation of the method is shown and numerical results based on the Euler-Lagrange equations will be presented in this dissertation.

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; 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 the above discussion can be found in [17].

2.1.1 Variational method of improving grid generation

Properties of grids such as length of grid lines or smoothness (I_s), area or volume of cells (I_v), and orthogonality of grid lines (I_o) are controlled by the minimization of a functional [18].

Brackbill and Saltzman in [19] combined these functionals into one functional:

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

where λ_v, λ_o are non-negative numbers.

It is shown by calculation in [19] that numerical solutions which minimize I is obtained for finite values of λ_o, λ_v.

Guojun Liao in [18] followed the approach of Roache and Steinberg but used a different smoothness functional I_s.

In [20], the functionals are:

$$I_s = \frac{1}{2}\iint_{\Omega}\left(|\nabla x|^2 + |\nabla y|^2\right)d\varsigma d\varphi,$$

$$I_v = \frac{1}{2}\iint_{\Omega}\left(\frac{D(x,y)}{D(\varsigma,\varphi)}\right)d\varsigma d\varphi,$$

$$I_o = \frac{1}{2}\iint_{\Omega}\left(x_\varsigma x_\varphi + y_\varsigma y_\varphi\right)^2 d\varsigma d\phi$$

where ∇x = the gradient of $x(\varsigma,\varphi)$, $\nabla x = \left(\frac{\partial x}{\partial \varsigma}, \frac{\partial x}{\partial \varphi}\right)$, $\nabla y = \left(\frac{\partial y}{\partial \varsigma}, \frac{\partial y}{\partial \varphi}\right)$ and

$\frac{D(x,y)}{D(\varsigma,\varphi)} = \det\begin{bmatrix} \frac{\partial x}{\partial \varsigma} & \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial \varsigma} & \frac{\partial y}{\partial \varphi} \end{bmatrix}$ is the Jacobian determinant.

The goal is to provide a trade-off between grid smoothness, orthogonality and uniform cell size. In [21], 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. In this thesis, the concern of grid folding is not an issue since we enforce the Jacobian determinant of the grid generated to be strictly positive through the control of a monitor function f.

J.U. Brackbill et al. in [19] quoted [22], I_v and I_o cannot be minimized separately because they do not have unique solutions; an example is used to prove this.

In [23], it is proven that the smoothness functional (I_s) has a unique solution. Based on this fact, the following cost functional I is minimized:

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

But it has not been proven that I has a solution.

Numerical solutions which minimize I are obtained for certain finite values of λ_o and λ_v.

In [18] a simpler model in 2D was studied theoretically. 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 is proven by the direct method in the calculus of variations. A regularity of minimum and a derivation that explains why I_V is used to control cell volume is also discussed.

Liao in [24] proved that there exists a $C^{0,\alpha}$ regularity of minimum of a functional related to the 3D grid-generation problem. The functional discussed is a linear combination of the length functional, the volume functional and a higher order smoothness term.

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

The direct method in the calculus of variations is used when the problem is to minimize an integral functional. The steps usually involve:

(i) Constructing a minimizing sequence.

(ii) Extracting a subsequence that converges to a limit weakly.

(iii) Show that the functional is weakly lower semi-continuous. Hence the subsequence converges strongly to the limit, which is a solution.

In this dissertation, we minimize I_o with I_s that is, we minimize the orthogonal functional with a penalty term to make the problem more elliptic.

We then derive the Euler-Lagrange equations and solve for u numerically.

2.1.2 Formulation of the Method

Minimize: $\int_D \left(\varphi_{x_1} \cdot \varphi_{x_2}\right)^2$

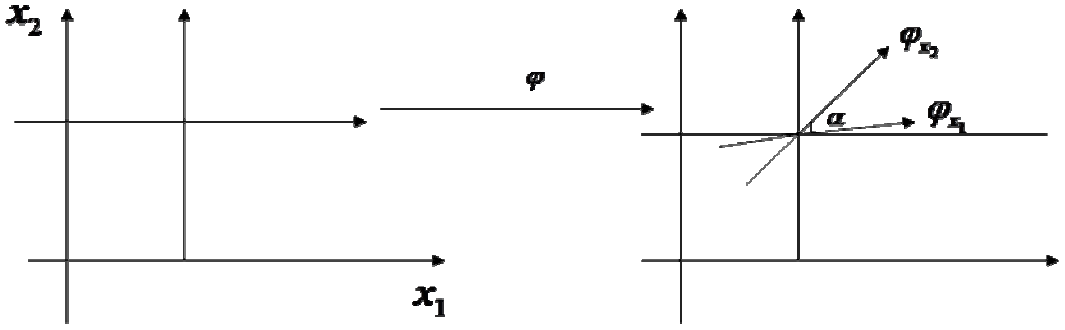


Figure 2.1 Illustration of orthogonality goal

Let D be a square domain in $\Re^2$ and $\varphi = (\varphi_1, \varphi_2)$. Define the cost functional I as

$$I := \frac{1}{2}\int_D \left(\varphi_{x_1} \cdot \varphi_{x_2}\right)^2 + \frac{\lambda}{2}\int_D |\nabla \varphi|^2 \qquad (2.1.2.1)$$

where $\lambda > 0$ and φ_{x_i} denotes derivatives of φ with respect to x_i for i = 1, 2 and

$$|\nabla \varphi|^2 = \varphi_{1x_1}^2 + \varphi_{1x_2}^2 + \varphi_{2x_1}^2 + \varphi_{2x_2}^2$$

So,

$$I = \frac{1}{2}\int_D \left(\varphi_{1x_1}\varphi_{1x_2} + \varphi_{2x_1}\varphi_{2x_2}\right)^2 + \frac{\lambda}{2}\int_D \left(\varphi_{1x_1}^2 + \varphi_{1x_2}^2 + \varphi_{2x_1}^2 + \varphi_{2x_2}^2\right) \qquad (2.1.2.2)$$

Now, let us introduce a small variation of φ; namely,

$$\varphi \quad \begin{cases} \varphi_1 \to \varphi_1 + \varepsilon\delta\varphi_1 \\ \varphi_2 \to \varphi_2 + \varepsilon\delta\varphi_2 \end{cases}$$

Now, let us take the Fréchet derivative of the cost function (2.1.2.2) with respect to φ_1 and φ_2, we obtain:

$$\frac{\partial I(\varphi+\varepsilon\delta\varphi)}{\partial\varphi_1} = I_{\varphi_1} = \frac{d}{d\varepsilon}\Big|_{\varepsilon=0}\frac{1}{2}\int_D\left[\left(\varphi_{1x_1}+\varepsilon(\delta\varphi_1)_{x_1}\right)\left(\varphi_{1x_2}+\varepsilon(\delta\varphi_1)_{x_2}\right)+\varphi_{2x_1}\varphi_{2x_2}\right]^2+\frac{\lambda}{2}\int_D\left(\left(\varphi_{1x_1}+\varepsilon(\delta\varphi_1)_{x_1}\right)^2+\left(\varphi_{1x_2}+\varepsilon(\delta\varphi_1)_{x_2}\right)^2+\varphi_{2x_1}^2+\varphi_{2x_2}^2\right)$$

$$\frac{\partial I(\varphi+\varepsilon\delta\varphi}{\partial\varphi_1} = I_{\varphi_1} = \frac{d}{d\varepsilon}\Big|\varepsilon=0\frac{1}{2}\int_D\left(\varphi_{1x_1}\varphi_{1x_2}+\varepsilon\varphi_{1x_1}(\delta\varphi_1)_{x_2}+\varepsilon\varphi_{1x_2}(\delta\varphi_1)_{x_1}+\varepsilon^2(\delta\varphi_1)_{x_1}(\delta\varphi_1)_{x_2}+\varphi_{2x_1}\varphi_{2x_2}\right)^2+\frac{\lambda}{2}\int_D\left(\left(\varphi_{1x_1}+\varepsilon(\delta\varphi_1)_{x_1}\right)^2+\left(\varphi_{1x_2}+\varepsilon(\delta\varphi_1)_{x_2}\right)^2+\varphi_{2x_1}^2+\varphi_{2x_2}^2\right)$$

$$= \int_D\left(\left(\varphi_{1x_1}(\delta\varphi_1)_{x_2}+\varphi_{1x_2}(\delta\varphi_1)_{x_1}\right)\left(\varphi_{1x_1}\varphi_{1x_2}+\varphi_{2x_1}\varphi_{2x_2}\right)\right)+\lambda\int_D\left((\delta\varphi_1)_{x_1}(\varphi_{1x_1})+(\delta\varphi_1)_{x_2}(\varphi_{1x_2})\right)$$

$$= \int_D\left[\left(\varphi_{1x_1}(\delta\varphi_1)_{x_2}\right)\left(\varphi_{1x_1}\varphi_{1x_2}\right)+\left(\varphi_{1x_1}(\delta\varphi_1)_{x_2}\left(\varphi_{2x_1}\varphi_{2x_2}\right)\right)+\left(\varphi_{1x_2}(\delta\varphi_1)_{x_1}\left(\varphi_{1x_1}\varphi_{1x_2}\right)\right)+\left(\varphi_{1x_2}(\delta\varphi_1)_{x_1}\left(\varphi_{2x_1}\varphi_{2x_2}\right)\right)+\lambda(\delta\varphi_1)_{x_1}(\varphi_{1x_1})+\lambda(\delta\varphi_1)_{x_2}(\varphi_{1x_2})\right]$$

$$= \int_D\left(\varphi_{1x_1}\varphi_{1x_1}\varphi_{1x_2}+\varphi_{1x_1}\varphi_{2x_1}\varphi_{2x_2}+\lambda\varphi_{1x_2}\right)(\delta\varphi_1)_{x_2}+\int_D\left(\varphi_{1x_2}\varphi_{1x_1}\varphi_{1x_2}+\varphi_{1x_2}\varphi_{2x_1}\varphi_{2x_2}+\lambda\varphi_{1x_1}\right)(\delta\varphi_1)_{x_1}$$
$$= \int_D\left(\varphi_{1x_1}^2\varphi_{1x_2}+\varphi_{1x_1}\varphi_{2x_1}\varphi_{2x_2}+\lambda\varphi_{1x_2}\right)(\delta\varphi_1)_{x_2}+\int_D\left(\varphi_{1x_2}^2\varphi_{1x_1}+\varphi_{1x_2}\varphi_{2x_1}\varphi_{2x_2}+\lambda\varphi_{1x_1}\right)(\delta\varphi_1)_{x_1}$$

In a similar way we derive I_{φ_2}

$$\frac{\partial I(\varphi+\varepsilon\delta\varphi}{\partial\varphi_2} = I_{\varphi_2} = \frac{d}{d\varepsilon}\Big|_{\varepsilon=0}\frac{1}{2}\int_D\left(\varphi_{1x_1}\varphi_{1x_2}+\left(\varphi_{2x_1}+\varepsilon(\delta\varphi_2)_{x_1}\right)\left(\varphi_{2x_2}+\varepsilon(\delta\varphi_2)_{x_2}\right)\right)^2+\frac{\lambda}{2}\int_D\left(\varphi_{1x_1}^2+\varphi_{1x_2}^2+\left(\varphi_{2x_1}+\varepsilon(\delta\varphi_2)_{x_1}\right)^2+\left(\varphi_{2x_2}+\varepsilon(\delta\varphi_2)_{x_2}\right)^2\right)$$

$$\frac{\partial I(\varphi+\varepsilon\delta\varphi}{\partial\varphi_2} = I_{\varphi_2} = \frac{d}{d\varepsilon}\Big|_{\mu=0}\frac{1}{2}\int_D\left(\varphi_{1x_1}\varphi_{1x_2}+\varphi_{2x_1}\varphi_{2x_2}+\varphi_{2x_1}\varepsilon(\delta\varphi_2)_{x_2}+\varphi_{2x_2}\varepsilon(\delta\varphi_2)_{x_1}\right)^2+\frac{\lambda}{2}\int_D\left(\varphi_{1x_1}^2+\varphi_{1x_2}^2+\left(\varphi_{2x_1}+\varepsilon(\delta\varphi_2)_{x_1}\right)^2+\left(\varphi_{2x_2}+\varepsilon(\delta\varphi_2)_{x_2}\right)^2\right)$$

$$= \int_D\left[\left(\varphi_{2x_1}(\delta\varphi_2)_{x_2}+\left(\varphi_{2x_2}(\delta\varphi_2)_{x_1}\right)\right)\left(\varphi_{1x_1}\varphi_{1x_2}+\varphi_{2x_1}\varphi_{2x_2}\right)+\lambda\int_D\left((\delta\varphi_2)_{x_1}\varphi_{2x_1}+(\delta\varphi_2)_{x_2}\varphi_{2x_2}\right)\right.$$

$$= \int_D\left(\varphi_{2x_1}(\delta\varphi_2)_{x_2}\varphi_{1x_1}\varphi_{1x_2}+\varphi_{2x_1}(\delta\varphi_2)_{x_2}\varphi_{2x_1}\varphi_{2x_2}+\varphi_{2x_2}(\delta\varphi_2)_{x_1}\varphi_{1x_1}\varphi_{1x_2}+\varphi_{2x_2}(\delta\varphi_2)_{x_1}\varphi_{2x_1}\varphi_{2x_2}+\lambda\varphi_{2x_1}(\delta\varphi_2)_{x_1}+\lambda\varphi_{2x_2}(\delta\varphi_2)_{x_2}\right)$$
$$= \int_D\left(\varphi_{2x_1}\varphi_{1x_1}\varphi_{1x_2}+\varphi_{2x_1}^2\varphi_{2x_2}+\lambda\varphi_{2x_2}\right)(\delta\varphi_2)_{x_2}+\int_D\left(\varphi_{2x_2}\varphi_{1x_1}\varphi_{1x_2}+\varphi_{2x_2}^2\varphi_{2x_1}+\lambda\varphi_{2x_1}\right)(\delta\varphi)_{x_1}$$

For ease of the notation, we define the following variables:

$$A := \varphi_{1x_2}^2 \varphi_{1x_1} + \varphi_{1x_2} \varphi_{2x_1} \varphi_{2x_2} \quad (2.1.2.3)$$

$$B := \varphi_{1x_1}^2 \varphi_{1x_2} + \varphi_{1x_1} \varphi_{2x_1} \varphi_{2x_2} \quad (2.1.2.4)$$

$$C := \varphi_{2x_2} \varphi_{1x_1} \varphi_{1x_2} + \varphi_{2x_2}^2 \varphi_{2x_1} \quad (2.1.2.5)$$

$$D := \varphi_{2x_1}^2 \varphi_{2x_2} + \varphi_{2x_1} \varphi_{1x_1} \varphi_{1x_2} \quad (2.1.2.6)$$

$$\tilde{A} := \varphi_{1x_2}^2 \varphi_{1x_1} + \varphi_{1x_2} \varphi_{2x_1} \varphi_{2x_2} + \lambda \varphi_{1x_1} \quad (2.1.2.7)$$

$$\tilde{B} := \varphi_{1x_1}^2 \varphi_{1x_2} + \varphi_{1x_1} \varphi_{2x_1} \varphi_{2x_2} + \lambda \varphi_{1x_2} \quad (2.1.2.8)$$

$$\tilde{C} := \varphi_{2x_2} \varphi_{1x_1} \varphi_{1x_2} + \varphi_{2x_2}^2 \varphi_{2x_1} + \lambda \varphi_{2x_1} \quad (2.1.2.9)$$

$$\tilde{D} := \varphi_{2x_1} \varphi_{1x_1} \varphi_{1x_2} + \varphi_{2x_1}^2 \varphi_{2x_2} + \lambda \varphi_{2x_2} \quad (2.1.2.10)$$

Then $I_{\varphi_1} = 0$ can be rewritten as: $\int_D \left(\tilde{A}, \tilde{B}\right) \cdot (\nabla \delta\varphi_1) = 0$ .Similarly, $I_{\varphi_2} = 0$ can be rewritten as $\int_D \left(\tilde{C}, \tilde{D}\right) \cdot (\nabla \delta\varphi_1) = 0$.

Applying the divergence theorem which says:

$$\int_{\partial\Omega} div\, F = \int_{\partial\Omega} F \cdot n \qquad (=0)$$

We let $F = hv$, $\quad div(hv) = h\,div v + \nabla h \cdot v$

where,

$h = \delta\varphi_1$

$v = \left(\tilde{A}, \tilde{B}\right)$

$\delta\varphi_1 = 0 \;\; on \;\; \partial\Omega$

$\Rightarrow I_{\varphi_1} = -\int_\Omega \left(div\left(\tilde{A}, \tilde{B}\right)\right)(\delta\varphi_1) = 0$

Since $\delta\varphi_1$ is arbitrary then $-div\left(\tilde{A}, \tilde{B}\right) = 0$

we get $-\int_D \left(\nabla \cdot \left(\tilde{A}, \tilde{B}\right)\right)\delta\varphi_1 = 0 \qquad \left(\delta\varphi_1 = 0 \ \ on \ \ \partial D\right)$

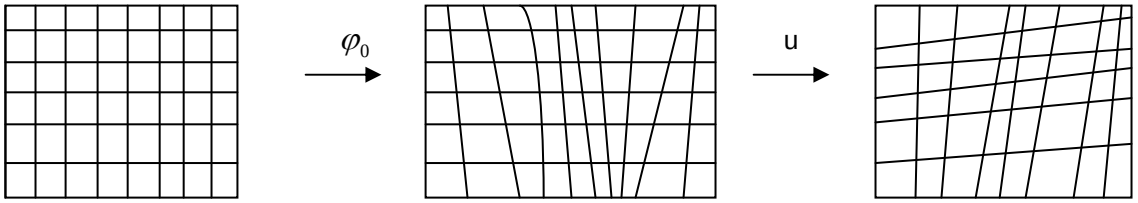
$\Rightarrow \ \nabla \cdot \left(\tilde{A}, \tilde{B}\right) = 0$ since $\delta\varphi_1$ is arbitrary.

Similarly, we have $\nabla \cdot \left(\tilde{C}, \tilde{D}\right) = 0$

The Euler-Lagrange equations are:

$\Delta\varphi_1$ = - div (A, B) / λ (2.1.2.11)

$\Delta\varphi_2$ = - div (C, D) / λ (2.1.2.12)



(a) (b) (c)

Figure 2.2 Uniform grid (a), Initial adaptive grid (b) and calculated grid (c).

Where,

$\varphi(x) = \varphi_o(x) + u(x)$, u(x) is a displacement to be added to the initial grid $\varphi_o(x)$.

The initial grid $\varphi_o(x)$ is generated by the grid deformation method discussed in chapter 1.

We implement on a 25 by 25 grid with $n \cdot u = 0$ on ∂D

Now,

Substitute $\varphi(x) = \varphi_o(x) + u(x)$ into (2.1.2.11) and (2.1.2.12) respectively, to get:

$\Delta\varphi(x) = \Delta\varphi_o(x) + \Delta u_1(x)$ = -div (A, B)/ λ (2.1.2.13)

$\Delta\varphi(x) = \Delta\varphi_o(x) + \Delta u_2(x)$ = -div (C, D)/ λ (2.1.2.14)

We find u_1 and u_2 from:

$$\Delta u_1(x) = \frac{-div(A,B)}{\lambda} - \Delta\varphi_{01}(x) \quad (2.1.2.15)$$

$$\Delta u_2(x) = \frac{-div(C,D)}{\lambda} - \Delta\varphi_{02}(x) \quad (2.1.2.16)$$

$$\Delta u_1(x) = \frac{-A_{x_1} - B_{x_2}}{\lambda} - \Delta\varphi_{01}(x) \quad (2.1.2.17)$$

$$\Delta u_2(x) = \frac{-C_{x_1} - D_{x_2}}{\lambda} - \Delta\varphi_{02}(x) \quad (2.1.2.18)$$

$$\Delta u_1(x) = \frac{-\left(\varphi_{1x_2}^2\varphi_{1x_1} + \varphi_{1x_2}\varphi_{2x_1}\varphi_{2x_2}\right)_{x_1} - \left(\varphi_{1x_1}^2\varphi_{1x_2} + \varphi_{1x_1}\varphi_{2x_1}\varphi_{2x_2}\right)_{x_2}}{\lambda} - \Delta\varphi_{01}(x) \quad (2.1.2.19)$$

$$\Delta u_2(x) = \frac{-\left(\varphi_{2x_2}^2\varphi_{1x_1}\varphi_{1x_2} + \varphi_{2x_2}^2\varphi_{2x_1}\right)_{x_1} - \left(\varphi_{2x_1}^2\varphi_{2x_2} + \varphi_{2x_1}\varphi_{1x_1}\varphi_{1x_2}\right)_{x_2}}{\lambda} - \Delta\varphi_{02}(x) \quad (2.1.2.20)$$

2.2 Numerical Scheme

We obtain the initial grid φ_0 , by the deformation method discussed in chapter 1.When we solve the Euler-Lagrange equations (2.1.2.19) and (2.1.2.20) we get the deformation of grids. The composition grids $\varphi_o + \mu u$ produce the significant improvement of orthogonality of grids.

In the numerical scheme, the initial grid φ_o is only needed once.

We set the Neumann boundary condition to 0, i.e. $n \cdot u = 0$. Computation of φ_o was done by calling a Laplace solver in FORTRAN. To solve the right hand side of (2.1.2.17) and (2.1.2.18), we call a poisson solver. When solving SOR we introduce a relaxation factor (real 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 (2.1.2.17) and (2.1.2.18) i.e. to show that our equation is correct. The computational residue gets closer to the tolerance, so 0.001 is sufficient.

Two types of boundaries are used: the dirichlet boundary and the neumann boundary, because initially we don't know the correct one to use.

2.2.1 Dirichlet boundary condition

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

2.2.2 Neumann boundary condition

This boundary condition imposes the imaginary grid points φ_o and the variation u near the boundary according to symmetric relation.

2.3 Numerical experimentation

We present several numerical results showing the improvement in orthogonality of the grid by first producing an initial grid φ_0 by the deformation method and then composing with u, which is found by solving the Euler-Lagrange Equations, and then the composition $\varphi_0 + \mu u$ produces the orthogonal grid. In numerical experiments, different multiples of u are added to φ_0, i.e. $\varphi = \varphi_0 + \mu u$.

The initial grid φ_0 is a 25 by 25 grid.

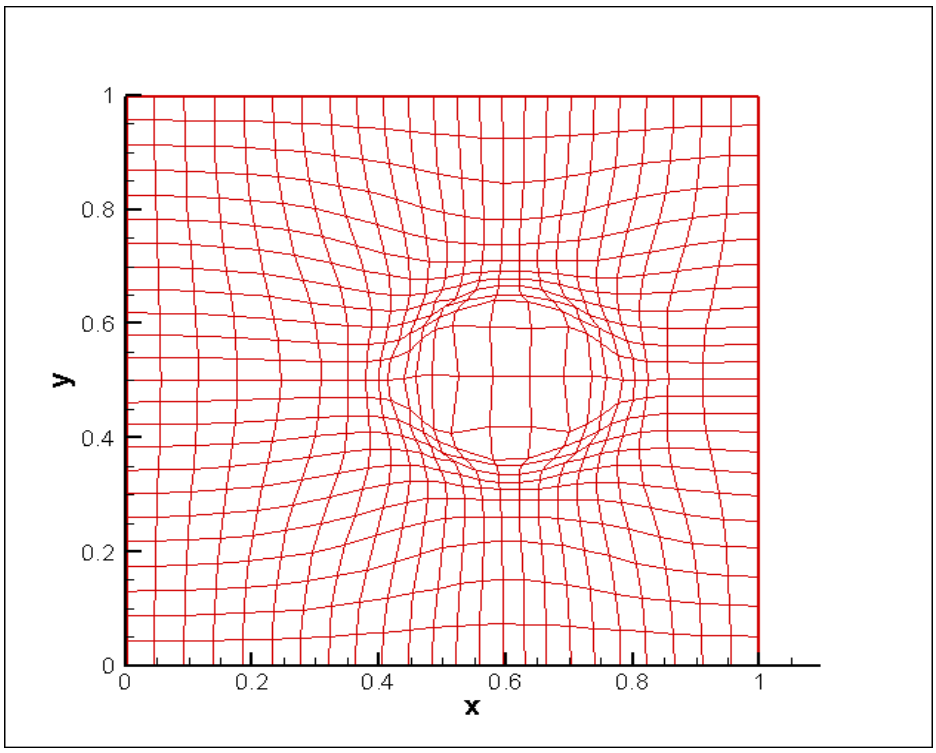
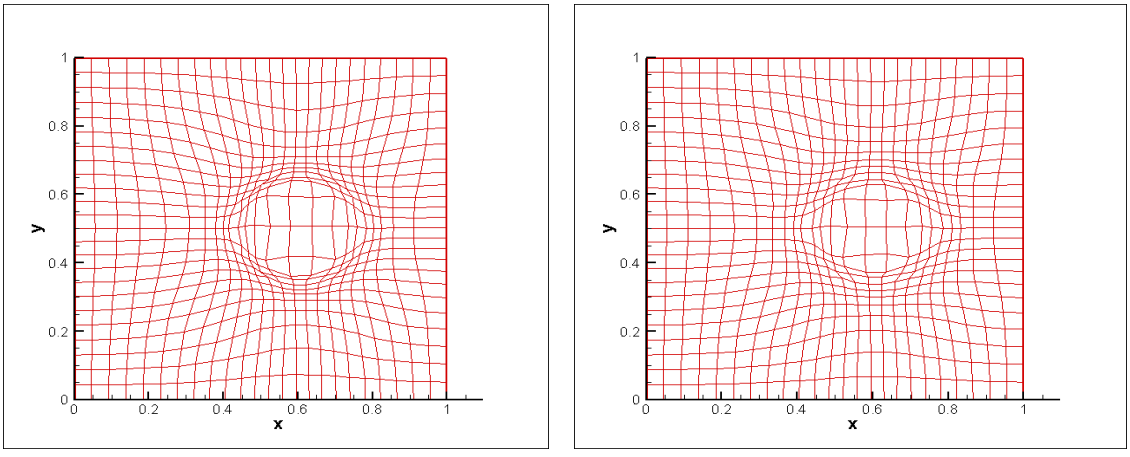


Figure 2.3 Initial Grid, φ_0

CASE 1: The control parameters are as follows:

Table 2.1 λ =1, μ = 0.2

Parameters	Value	Explanation
λ	1	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	1	boundary type:1-Dirichlet condition; 2-Neumann condition
bdfactor	0	bdfactor: 1-do computation on boundary, 0-not on boundary
μ	0.2	the quantity level for adding the change u



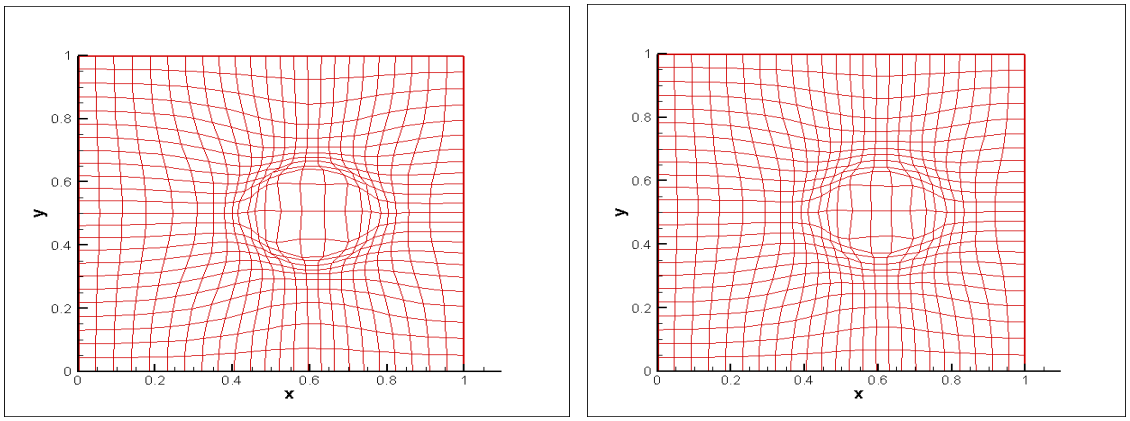
(a) (b)

Figure 2.4 Comparison between Initial adaptive grid, φ_0 (a) and the calculated grid (b)

CASE 2: The control parameters are as follows:

Table 2.2 λ =0.9, μ = 0.2

Parameters	Value	Explanation
λ	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 condition; 2-Neumann condition
bdfactor	0	bdfactor: 1-do computation on boundary, 0-not on boundary
μ	0.2	the quantity level for adding the change u



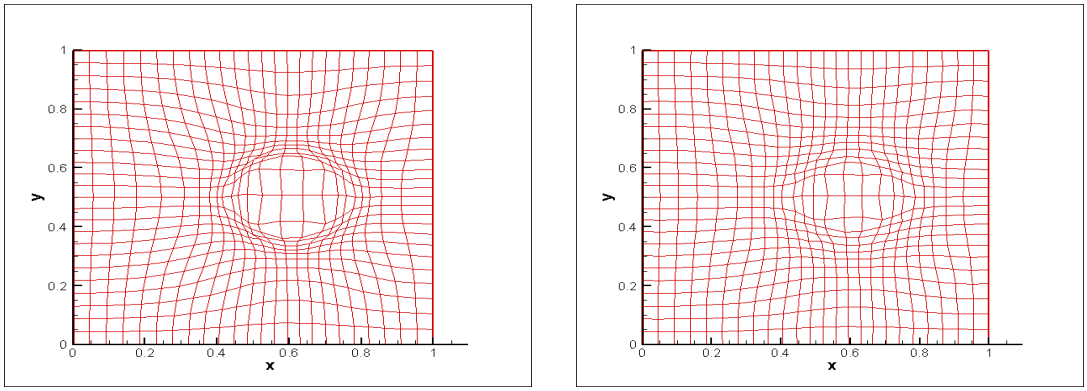
(a) (b)

Figure 2.5 Comparison between Initial adaptive grid, φ_0 (a) and the calculated grid (b)

CASE 3: The control parameters are as follows:

Table 2.3 λ =0.7, μ = 0.4

Parameters	Value	Explanation
λ	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 condition; 2-Neumann condition
bdfactor	0	bdfactor: 1-do computation on boundary, 0-not on boundary
μ	0.4	the quantity level for adding the change u



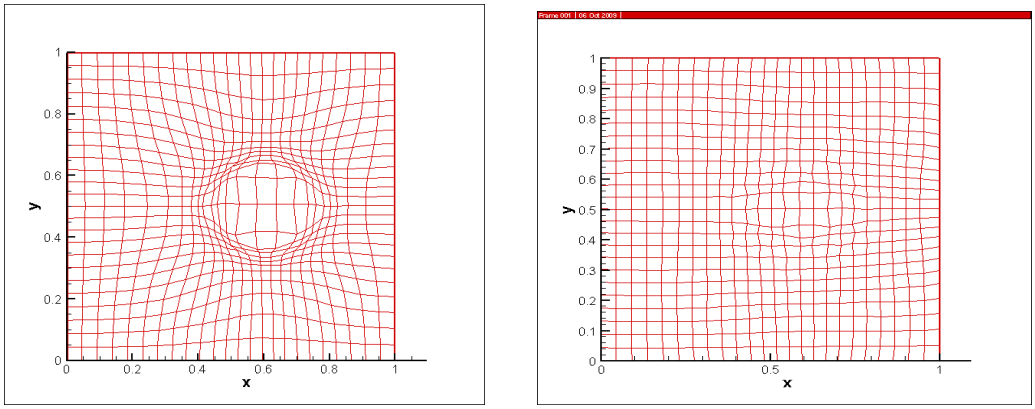
(a) (b)

Figure 2.6 Comparison between Initial adaptive grid, φ_0 (a) and the calculated grid (b)

CASE 4: The control parameters are as follows:

Table 2.4 λ =0.7, μ = 0.9

Parameters	Value	Explanation
λ	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 condition; 2-Neumann condition
bdfactor	0	bdfactor: 1-do computation on boundary, 0-not on boundary
μ	0.9	the quantity level for adding the change u



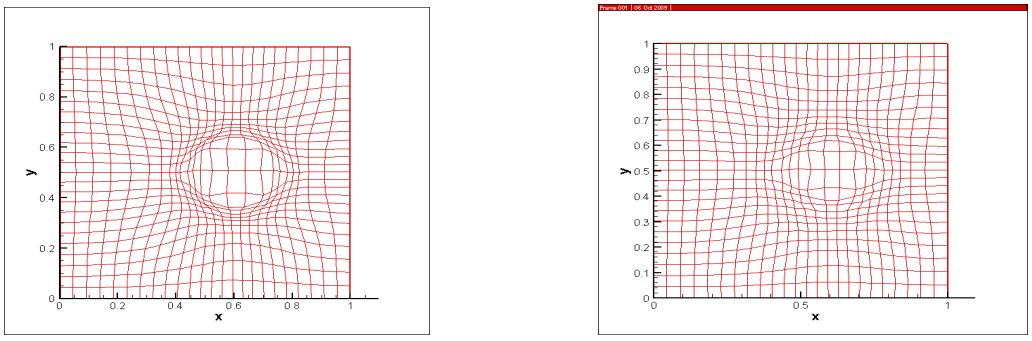
(a) (b)

Figure 2.7 Comparison between Initial adaptive grid, φ_0 (a) and calculated grid (b)

CASE 5: The control parameters are as follows:

Table 2.5 λ =0.69, μ = 0.4

Parameters	Value	Explanation
λ	0.69	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 condition; 2-Neumann condition
bdfactor	0	bdfactor: 1-do computation on boundary, 0-not on boundary
μ	0.4	the quantity level for adding the change u



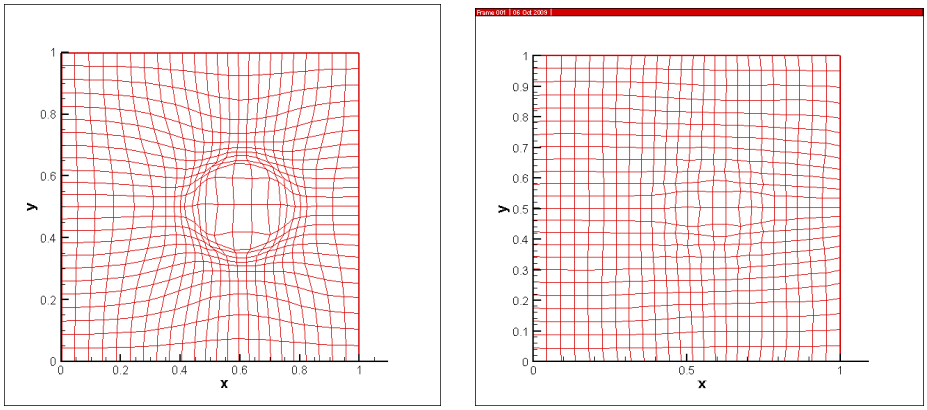
(a) (b)

Figure 2.8 Comparison between Initial adaptive grid, φ_0 (a) and the calculated grid (b)

CASE 6: The control parameters are as follows:

Table 2.6 λ =0.69, μ = 0.9

Parameters	Value	Explanation
λ	0.69	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 condition; 2-Neumann condition
bdfactor	0	bdfactor: 1-do computation on boundary, 0-not on boundary
μ	0.9	the quantity level for adding the change u



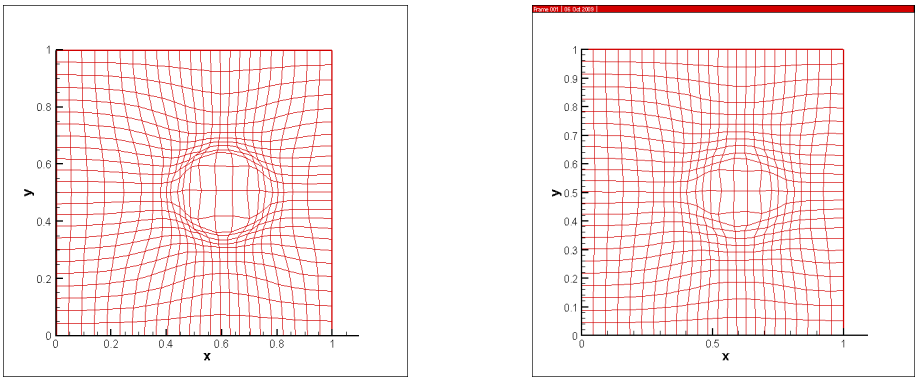
(a) (b)

Figure 2.9 Comparison between Initial adaptive grid, φ_0 (a) and the calculated grid (b)

CASE 7: The control parameters are as follows:

Table 2.7 λ =0.75, μ = 0.4

Parameters	Value	Explanation
λ	0.75	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 condition; 2-Neumann condition
bdfactor	0	bdfactor: 1-do computation on boundary, 0-not on boundary
μ	0.4	the quantity level for adding the change u



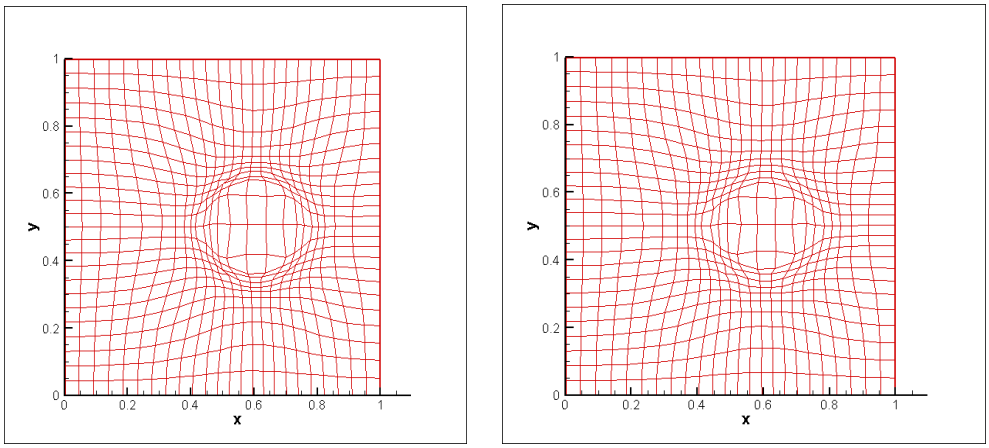
(a) (b)

Figure 2.10 Comparison between Initial adaptive grid, φ_0 (a) and the calculated grid (b)

CASE 8: The control parameters are as follows:

Table 2.8 λ =50, μ = 0.2

Parameters	Value	Explanation
λ	50	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 condition; 2-Neumann condition
bdfactor	0	bdfactor: 1-do computation on boundary, 0-not on boundary
μ	0.2	the quantity level for adding the change u



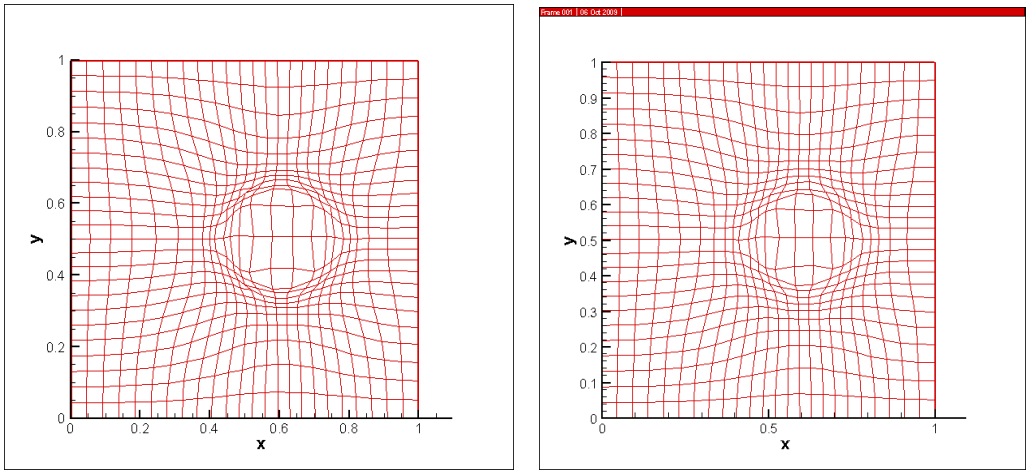
(a) (b)

Figure 2.11 Comparison between Initial adaptive grid, φ_0 (a) and the calculated grid (b)

CASE 9: The control parameters are as follows

Table 2.9 λ =100, μ = 0.2

Parameters	Value	Explanation
λ	100	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 condition; 2-Neumann condition
bdfactor	0	bdfactor: 1-do computation on boundary, 0-not on boundary
μ	0.2	the quantity level for adding the change u



(a) (b)

Figure 2.12 Comparison between Initial adaptive grid, φ_0 (a) and the calculated grid (b)

2.4 Discussion

In case 1, with λ=1 the computation converges after 7 iteration. Slight orthogonality is noticed [Figure 2.4].

In case 2 with λ=0.9, the computation converges after 6 iterations, more orthogonality is noticed [Figure 2.5].

In case 3 with λ=0.7, the computation converges after 11 iterations, significant orthogonal grid noticed [Figure 2.6].

In case 6 with λ=0.69, the computation converges after 12 iterations and significantly better orthogonality is noticed [figure 2.9].

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

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

We discovered that for higher values of λ, computation converges after 17 iterations.

In order to avoid distortion of grids {instability}, λ has to be between $0.6 \le \lambda \le 1.0$ and μ between $0 \le \mu \le 1$.

To conclude, the smaller λ is and the larger μ is, the nearly orthogonal the grid.

With this method we do not have to worry about the grid folding as some techniques did, 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 do not have to add the volume functional, I_v to generate good results.

CHAPTER 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: changing 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.

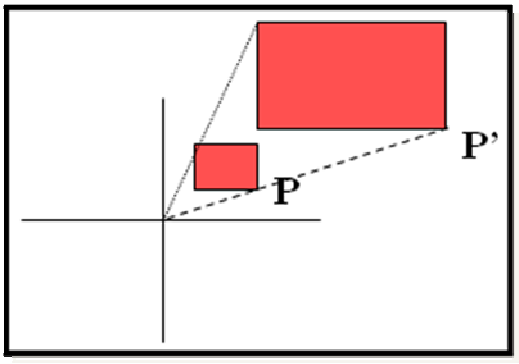


Figure 3.1 2D scaling from the origin

Let point P be defined as $P(x, y)$.

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.

$x' = s_x \cdot x, \qquad y' = s_y \cdot y$

Define the matrix, $S = \begin{bmatrix} s_x & 0 \\ 0 & s_y \end{bmatrix}$

Now,

$$P' = SP \text{ or } \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} s_x & 0 \\ 0 & s_y \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

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 dissertation, our problem concerns 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 [11].

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 [25].

Definition 3.2.1: Let $d \in \aleph$ and $R, T \in \operatorname{Im} g(d)$. The sum of squared differences (SSD) distance measure D^{SSD} is defined by $D^{SSD} : \operatorname{Im} g(d)^2 \to \Re$,

$$D^{SSD}[R,T] := \frac{1}{2}\|T - R\|_{L_2}^2 = \frac{1}{2}\int_{\Re^d} \left(T(x) - R(x)\right)^2 dx$$

For a transformation $\Phi : \Re^d \to \Re^d$ we also define:

$$D^{SSD}[R,T;\Phi] = D^{SSD}[R,T \circ \Phi] = \frac{1}{2}\int_{\Re^d} \left(T(\Phi(x)) - R(x)\right)^2 dx$$

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 [26] and Collignon et al in [27] and has been used since by Kim et al in [28] 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 [11].

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 were discussed in chapter 1 and chapter 4.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 [29].

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 $J[u] := D[R, T; u] + \alpha S[u]$ 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 [11].

CHAPTER 4

OPTIMAL CONTROL APPROACH OF AN IMAGE REGISTRATION PROBLEM

In this chapter we will describe the optimal control approach to non-rigid image registration based on the sum of squared differences (SSD) using the grid deformation method as constraints. We will then prove the existence of optimal solutions, existence of Lagrange multipliers using a div-curl constraint.

Finally, we will then derive the optimality system from which optimal solutions can be calculated. An ODE constraint is used in [30] to find Φ, the time-dependent mapping.

The existence of solution is proved using an existence theorem and two well-known results in functional analysis.

4.1 Grid Deformation Method (Used in Image Registration)

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 [31], which is essential in constructing differentiable and invertible transformations to solve mesh adaptation problems.

The deformation method has its origin in differential geometry [32]. 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. With this unwanted registration results are avoided.

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.

For image registration, we use the grid deformation equations as constraints.

4.2 Set-up of the Cost (Objective) Functional in 2D

Let $\Omega \subset \Re^2$ be a bounded domain that is a convex polygon or that has a $C^{1,1}$ boundary $(\partial\Omega)$.

Given:

R(x): Reference image fixed

S(x): a differentiable and invertible transformation from $\Omega \to \Omega$

We assume that S~ 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 $R(x) \sim T(\Phi(x))$ i.e. R(x) is as close to $T(\Phi(x))$ as possible in L^2 norm.

This means in a practical sense that:

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

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

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

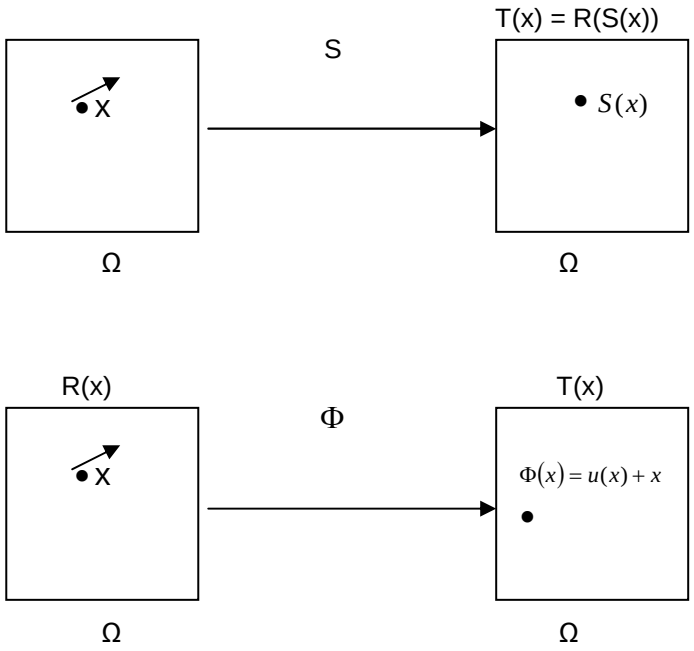


Figure 4.1 Illustration of registration goal

Where, u(x) is a displacement vector field,

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

$$x \in \Omega$$

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

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

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

The Sobolev spaces are defined as follows:

$$H^m(\Omega) = \left\{p \in L^2(\Omega) : D^\alpha p \in L^2(\Omega) \text{ for } 0 \le |\alpha| \le m\right\}, \quad m = 1, 2.$$

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 $\|.\|_m, \langle .,. \rangle_m$ denote the corresponding norm and inner product respectively.

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

We define the functional:

$$J(u,f,g)=\frac{1}{2}\|T(\Phi(x))-R(x)\|_{L^2}{}^2+\frac{d_1}{2}\|f\|_{H^1}{}^2+\frac{d_2}{2}\|g\|_{H^1}{}^2 \qquad (4.2.1)$$

where, $d_1>0$ and $d_2>0$ are penalty parameters to be chosen for numerical stability.

Then we minimize $J(u,f,g)$ with respect to the scalar control functions f and g, subject to the following constraints:

$$div\,u=u_{1x_1}+u_{2x_2}=f-1 \qquad in \quad \Omega$$
$$curl\,u=u_{2x_1}-u_{1x_2}=g \qquad in \quad \Omega.$$

Note 1: The displacement vector field u can not be determined from one scalar equation alone, additional equations are needed. According to the Helmholtz Decomposition Theorem any smooth, rapidly decaying vector field (u in our case) can be decomposed into the sum of an irrotational (curl-free) vector field and a solenoidal (divergence free) vector field. Thus, it is natural to add the above curl equation.

Note 2: The last two terms in (4.2.1) are needed to prevent the controls from becoming unbounded.

Note 3: According to [30] using a L^2 - norm penalization of f and g, 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. Therefore, a stronger norm, the H^1 norm is used in (4.2.1).

The following theorem explains why the right hand side of $divu$ is $f-1$.

Theorem 4.2.1: Linearization of Φ at identity [9]

Let $\Omega \subset \Re^2$ be a domain. For $\Phi(x) = \varepsilon u(x) + x = \left(\varepsilon u_1\left(x_1, x_2\right) + x_1, \varepsilon u_2\left(x_1, x_2\right) + x_2\right)$. Then Jacobian determinant of Φ, $I\left(\Phi\right)$ satisfies $I\left(\Phi\right) = 1 + div\left(\varepsilon u\right) + \mathrm{O}\left(\varepsilon^2\right)$ for every $\varepsilon > 0$ in Ω.

PROOF:

$$I\left(\Phi\right) = \begin{vmatrix} \dfrac{\partial u_1}{\partial x_1} & \dfrac{\partial u_1}{\partial x_2} \\ \dfrac{\partial u_2}{\partial x_1} & \dfrac{\partial u_2}{\partial x_2} \end{vmatrix} = \begin{vmatrix} 1 + \varepsilon u_{1x_1} & \varepsilon u_{1x_2} \\ \varepsilon u_{2x_1} & 1 + \varepsilon u_{2x_2} \end{vmatrix} =$$

$$= \left(1 + \varepsilon u_{1x_1}\right)\left(1 + \varepsilon u_{2x_2}\right) - \left(\varepsilon u_{1x_2}\right)\left(\varepsilon u_{2x_1}\right) = 1 + \varepsilon u_{2x_2} + \varepsilon u_{1x_1} + \varepsilon^2 u_{1x_1} u_{2x_2} - \varepsilon^2 u_{1x_2} u_{2x_1}$$

=

$$1 + \varepsilon u_{2x_2} + \varepsilon u_{1x_1} + \varepsilon^2\left(u_{1x_1} u_{2x_2} - u_{1x_2} u_{2x_1}\right) = 1 + \varepsilon u_{2x_2} + \varepsilon u_{1x_1} + \mathrm{O}(\varepsilon^2) = 1 + div(\varepsilon u) + \mathrm{O}(\varepsilon^2).$$

□

Let us define the product Hilbert Spaces:

$$V = \left[H^2(\Omega)\right]^2 \times H^1(\Omega) \times H^1(\Omega)$$
$$W = H^1(\Omega) \times H^1(\Omega) \times H^{\frac{3}{2}}(\partial\Omega)$$

And define the constraint operator in a weak formulation:
$M : V \to W$ as follows:

$M(u, f, g) = 0$ is the constraint equation

$$\Leftrightarrow$$

$$\left.\begin{aligned} \langle \nabla \cdot u - f + 1, \xi \rangle &= 0 \qquad \forall \xi \in H^{-1}(\Omega) \\ \langle \nabla \times u - g, \eta \rangle &= 0 \qquad \forall \eta \in H^{-1}(\Omega) \\ \langle n \cdot u, \nu \rangle &= 0 \qquad \forall \nu \in H^{-\frac{3}{2}}(\partial\Omega) \end{aligned}\right\} \tag{4.2.2}$$

The admissible set of all solutions is defined by:

$$U_{ad} = \{(u, f, g) \in V : J(u, f, g) \text{ is bounded and } M(u, f, g) = 0\} \tag{4.2.3}$$

Then the optimal control problem is given by:

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

(4.2.4)

4.3 Existence of Optimal Solutions

We next show the existence of solutions of the optimal control problem (4.2.4).

Under the assumptions we have made about $\Omega \subset \Re^2$, it is well known that given f,g$\in H^1(\Omega)$,

$\exists\ a\quad u \in \left[H^2(\Omega)\right]^2$ satisfying:

$$\nabla \cdot u = f - 1 \quad and \quad \nabla \times u = g \quad in \quad \Omega, \quad n \cdot u = 0 \quad on \quad \partial\Omega$$

Since $u \in \left[H^2(\Omega)\right]^2$, by the Sobolev imbedding theory, u is bounded and Hölder-continuous

i.e. $\forall x, y \in \Omega$ for some $K_1 > 0$, $|u(x)| \le K_1$ and for some constant $K_2 > 0$,

$|u(x) - u(y)| \le K_2 |x - y|^{\lambda}, \quad for \quad 0 < \lambda < 1$

Also, the div-curl system described in (4.2.1) 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,

$$\begin{aligned} &div\ u = f - 1 \\ &curl\ u = g \qquad\qquad \text{is Elliptic in the Petrovski sense} \\ &Ru = 0 \qquad on \quad \partial\Omega \end{aligned}$$

PROOF:

Now look at the div-curl operator

$$N(u) = \begin{pmatrix} \nabla \cdot u \\ \nabla \times u \end{pmatrix} = \begin{pmatrix} f - 1 \\ g \end{pmatrix} \text{ in } \Omega$$

$$\text{with } \ n \cdot u = 0 \quad in \quad \partial\Omega$$

Let $f^* = f - 1$

M is elliptic in the sense of Petrovski (i.e. regular Ellipticity)

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

where,

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

Control variables: f^* and g {scalar valued functions} are in $H^1(\Omega)$

We can write the above as a linear system as:

$$= A\frac{\partial u}{\partial x_1} + B\frac{\partial u}{\partial x_2} = h \qquad \text{in } \Omega\,.$$

$$Ru = u \cdot n = 0 \quad in \quad \partial\Omega$$

or

$$Ru = u \times n = 0 \quad in \quad \partial\Omega$$

where,

$$h = \begin{pmatrix} f^* \\ g \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

n is a unit outward normal to Ω

R is a full rank 1×2 matrix

We express the operator as follows:

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

Note that ellipticity in the sense of Petrovski requires that:

$$\det(A\lambda + B\eta) = 0 \quad \Leftrightarrow \quad \lambda = 0 \ \ and \ \ \eta = 0 \ \text{ for } (\lambda, \eta) \in \Re^2$$

Now,

Assume $\lambda \neq 0 \ and \ \eta \neq 0$

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

□

As a consequence of the div-curl system being elliptic, there exists a constant

C > 0 such that:

$$\frac{1}{C}\|v\|_1 \le \|\nabla\cdot v\|_0 + \|\nabla\times v\|_0 \le C\|v\|_1$$

with $v\in H_n^1(\Omega)$.

Next, we prove the existence of optimal solutions in the following theorem.

THEOREM 4.3: There exists a solution $(u, f, g)\in U_{ad}$ for the optimal control problem (4.2.4)

PROOF:

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

$\Rightarrow\ u=0$

$\Rightarrow\ U_{ad}\neq\phi$ because $(0,1,0)\in U_{ad}$

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 (4.2.3), f^n *and* g^n are bounded in $H^1(\Omega)$, $\|f^n\|, \|g^n\|$ *are bounded* .

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

Then we can extract subsequences such that:

$$\begin{aligned} u^n &\to \tilde{u} \quad \textit{weakly in} \quad \left[H^2(\Omega)\right]^2 \\ f^n &\to \tilde{f} \quad \textit{weakly in} \quad H^1(\Omega) \\ g^n &\to \tilde{g} \quad \textit{weakly in} \quad H^1(\Omega) \end{aligned} \tag{4.2.5}$$

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

Next step is to show that the limit $(\tilde{u}, \tilde{f}, \tilde{g}) \in V$ satisfies the constraint equations in (4.2.2)

$\tilde{u}$ and $\tilde{f}$ satisfy $\langle \nabla \cdot u - f + 1, \xi \rangle = 0 \qquad \forall \xi \in H^{-1}(\Omega)$

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 (4.2.5),

$$\int_\Omega \left(\nabla \cdot u^n - f^n + 1 - \left(\nabla \cdot \tilde{u} - \tilde{f} + 1\right)\right) \cdot \xi \, d\Omega =$$

$$\int_\Omega \left(\nabla \cdot u^n - \nabla \cdot \tilde{u} - f^n + \tilde{f}\right) \cdot \xi \, d\Omega = \int_\Omega \left(\nabla \cdot u^n - \nabla \cdot \tilde{u} - \left(f^n - \tilde{f}\right)\right) \cdot \xi \to 0$$

Now let us show that:

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

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

By (4.2.5),

$$\int_\Omega \left(\nabla \times u^n - g^n - \left(\nabla \cdot \tilde{u} - \tilde{g} \right) \right) \cdot \eta \; d\Omega =$$
$$\int_\Omega \left(\left(\nabla \times u^n - \nabla \times \tilde{u} \right) - \left(g^n - \tilde{g} \right) \right) \cdot \eta \; d\Omega \to 0$$

Also, 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 (4.2.5).

Thus, $\left(\tilde{u}, \tilde{f}, \tilde{g} \right) \in V$ satisfies $M(u, f, g) = 0$.

Finally, since Ω is a convex polygon in $\Re^2$, its interior is a convex set, which is closed and bounded.

This implies that J (.,.) 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 (.,.) is weakly lower semi-continuous

i.e. $\lim_{(u^n, f^n, g^n) \to (\tilde{u}, \tilde{f}, \tilde{g})} J\left(u^n, f^n, g^n \right) \geq J\left(\tilde{u}, \tilde{f}, \tilde{g} \right)$

$\Rightarrow$

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

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) into an unconstrained problem and then to derive an optimality system. Thus, 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 (4.2.2). Let $M'(u, f, g) \in L(V;W)$ be the first Frechét derivative of M at (u, f, g) defined as $M'(u, f, g)\cdot(\tilde{u}, \tilde{f}, \tilde{g}) = F$ for $(\tilde{u}, \tilde{f}, \tilde{g}) \in V$ and $F = (f_1, f_2, f_3) \in W$

iff

$$\left.\begin{aligned} \langle \nabla\cdot\tilde{u} - \tilde{f}, \xi\rangle &= \langle f_1, \xi\rangle && \forall \xi \in H^{-1}(\Omega) \\ \langle \nabla\times\tilde{u} - \tilde{g}, \eta\rangle &= \langle f_2, \eta\rangle && \forall \eta \in H^{-1}(\Omega) \\ \langle n\cdot\tilde{u}, \nu\rangle &= \langle f_3, \nu\rangle && \forall \nu \in H^{-\frac{3}{2}}(\partial\Omega) \end{aligned}\right\} \qquad (4.5.1)$$

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 [36].

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 $\overset{\circ}{u}$ 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.) $F : nbhd(\overset{\circ}{u}) \subset V_1 \to \Re$ is Frechet-differentiable at $\hat{u}$;
(2.) G is continuously Frechet-differentiable at $\hat{u}$;
(3.) $G'(\overset{\circ}{u}) : V_1 \to V_2$ is onto

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

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

$\langle .,. \rangle$ denotes the duality pairing between V_2 and $(V_2)^*$ and $F'(\hat{u})v$ and $G'(\overset{\circ}{u})$v denote the actions of $F'(\hat{u})$ as an operator mapping $v \in V_1$ into V_2, respectively.

A proof can be found in [33], Theorem 43.19.

We fit our optimization problem into this abstract framework. F and G in the above theorem are our J and M, respectively.

To show the existence of Lagrange multipliers, we first prove that the operator $M'(u, f, g)$ is onto.

In the following lemma, we prove M' is onto.

Lemma 4.5.1: Let (u, f, g) be a solution to the optimal control problem given in (4.2.4). Then the operator $M'(u, f, g)$ is onto W.

PROOF:

Let $(f_1, f_2, f_3) \in W$. Since the system $M'(u, f, g)\cdot(\tilde{u}, \tilde{f}, \tilde{g}) = (f_1, f_2, f_3)$ is underdetermined i.e. there are more variables than equations, we choose some variables and then solve for the rest.

We choose any $\tilde{f} \in H^1(\Omega)$ and choose any $\tilde{g} \in H^1(\Omega)$.

For $f_3 \in H^{\frac{3}{2}}(\partial\Omega)$, by the trace theorem, there is a $v \in H^2(\Omega)$ such that $n \cdot v = f_3$.

Let $\tilde{u} = w + v$ then we show that $(\tilde{u}, \tilde{f}, \tilde{g})$ satisfies (4.5.1)

$$\nabla \cdot w = \tilde{f} + f_1 - \nabla \cdot v \quad in \quad \Omega$$
$$\nabla \times w = \tilde{g} + f_2 - \nabla \times v \quad in \quad \Omega$$
$$n \cdot w = f_3 \quad on \quad \partial\Omega$$

$$\nabla \cdot w = \nabla \cdot (\tilde{u} - v) = \nabla \tilde{u} - \nabla v = \tilde{f} + f_1 - \nabla \cdot v = \tilde{f} + f_1$$
$$\nabla \tilde{u} = \tilde{f} + f_1$$

$$\nabla \times w = \nabla \times (\tilde{u} - v) = \nabla \times \tilde{u} - \nabla \times v = \tilde{g} + f_2 - \nabla \times v$$
$$\nabla \times \tilde{u} = g + f_2$$

$$n \cdot (\tilde{u} - v) = f_3$$
$$n\tilde{u} - nv = f_3$$
$$n \cdot \tilde{u} = n \cdot w + n \cdot v = f_3$$
$$n \cdot \tilde{u} = f_3$$

We have found a $(\tilde{u}, \tilde{f}, \tilde{g})$ satisfying (4.5.1) □

Now, we will prove that a non-zero Lagrange multiplier exists by following Theorem 4.5.1

We construct the Lagrangian functional as follows:

$$L(u, f, g, \Theta_1, \Theta_2) = J(u, f, g) + \int_\Omega \Theta_1 (divu - (f - 1)) + \int_\Omega \Theta_2 (curlu - g)$$

and then taking the derivative we have:

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

where,

$$(M'(\hat{u})v) \cdot (v_1, v_2) = (divu - (f - 1), curlu - g) \cdot (v_1, v_2)$$

and Θ_1, Θ_2 *are Lagrange multilpiers.*

The above is used in Theorem 4.5.2 below to prove the existence of multipliers.

THEOREM 4.5.2: Let $(u, f, g) \in V$ denote an optimal solution to the optimal control problem previously discussed. Then $\exists$ a non-zero Lagrange multiplier $\Theta \in W^*$ such that:

$$J'(u, f, g) \cdot (\tilde{u}, \tilde{f}, \tilde{g}) + \langle M'(u, f, g) \cdot (\tilde{u}, \tilde{f}, \tilde{g}), \Theta \rangle = 0 \qquad \forall (\tilde{u}, \tilde{f}, \tilde{g}) \in V$$

PROOF:

According to [33], M' being onto implies the existence of the Lagrange Multipliers. For completeness we include the arguments in [33] here.

Consider the non-linear operator $N : V \to \Re \times W$ defined by:

$N(\tilde{u}, \tilde{f}, \tilde{g}) = \begin{pmatrix} J(\tilde{u}, \tilde{f}, \tilde{g}) - J(u, f, g) \\ M(\tilde{u}, \tilde{f}, \tilde{g}) \end{pmatrix}$, where J and M are defined as in (4.2.1) and (4.2.2), respectively.

Then the Frechét derivative $N'(u, f, g)$ is defined as:

$N'(u, f, g) \cdot (\tilde{u}, \tilde{f}, \tilde{g}) = (a, F)$ *for* $(\tilde{u}, \tilde{f}, \tilde{g}) \in V$ and $(a, F) \in \Re \times W$

$$\Leftrightarrow$$

$$\left\langle \nabla T(\Phi(x))(T(\Phi(x))-R(x)), \tilde{\Phi} \right\rangle + d_1 \left\langle f, \tilde{f} \right\rangle + d_2 \langle g, \tilde{g} \rangle + \omega_1 \left\langle \nabla f, \nabla \tilde{f} \right\rangle + \omega_2 \langle \nabla g, \nabla \tilde{g} \rangle = a$$

where, $$\Phi(x) = u(x) + id$$

$$M'(u, f, g) \cdot (\tilde{u}, \tilde{f}, \tilde{g}) = F$$

The operator $M'(u, f, g)$ is onto W by Lemma 4.5.1 and therefore has a closed range in W.

Also, we know that $M'(u, f, g)$ is a linear operator from V to W.

Thus, the kernel of $M'(u, f, g)$; $\ker(M'(u, f, g))$ is a closed subspace of V.

Since $J'(u, f, g)$ acting on the kernel of $M'(u, f, g)$ is either identically 0 or onto $\Re$ (this follows from the fact that whenever f is a linear functional on a Banach space X, then either $f \equiv 0$ or the range of f, Im(f) is $\Re$.

$J'(u, f, g)$ acting on the kernel of $M'(u, f, g)$ has a closed range.

Now, we recall the following well-known result:

For X, Y, Z Banach spaces, let $A: X \to Y$ and $B: X \to Z$ be linear continuous operators.

If the range of A is closed in Y and B (ker (A)) is closed in Z, then $C: X \to Y \times Z$ defined by:

Cx= (Ax, Bx) has a closed range in $Y \times Z$.

$N'(u, f, g)$ has a closed range in $\Re \times W$.

Now, we suppose that $N'(u, f, g)$ is onto.

Then $\exists$ a $(\tilde{u}, \tilde{f}, \tilde{g}) \in U_{ad}$ satisfying $J(\tilde{u}, \tilde{f}, \tilde{g}) < J(u, f, g)$,

where,

$J(u - \tilde{u}, f - \tilde{f}, g - \tilde{g}) < \varepsilon$, with a small $\varepsilon > 0$ which contradicts the hypothesis that (u, f, g) is an optimal solution to (4.2.4).

To conclude, we use the Hahn-Banach Theorem.

Since, $\mathrm{Im}\left(N'(u,f,g)\right)$ is closed in $\Re \times W$, for any $(a,F) \in \mathrm{Im}\left(N'(u,f,g)\right)$, $\exists$ a non-zero $\Theta \in (\Re \times W)^*$ 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 $\tilde{a} = -1$ and therefore the theorem holds.

So, the penalized Lagrangian functional is:

$$L(u,f,g,\xi,\eta,\nu) = J(u,f,g) + \int_\Omega (\nabla \cdot u - f + 1)\xi dx + \int_\Omega (\nabla \times u - g)\eta dx + \int_{\partial\Omega} (n \cdot u)\nu ds$$

$$(\xi,\eta,\nu) \in W^*$$

$$W^* = H^{-1}(\Omega) \times H^{-1}(\Omega) \times \Re \times H^{-\frac{3}{2}}(\partial\Omega) \qquad \square$$

In the next section we derive the optimality system from which optimal solutions can be calculated.

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 4.6: The equalities hold for any scalar function h defined on Ω and a vector v.

$$\int_\Omega \nabla \cdot (hv) = \int_\Omega \nabla h \cdot v + \int_\Omega h \nabla \cdot v$$

$$\int_\Omega \nabla \cdot (hv) = \int_{\partial\Omega} (hv) \cdot \vec{n} = 0, \ if \ \ h = 0 \ \ on \ \ \partial\Omega$$

Where $\vec{n}$ is the outward and unit normal vector.

State Equations: The state equations are obtained by solving the equations $L_\xi = 0$ and $L_\eta = 0$

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

$$= \int_\Omega (\nabla \cdot u - f + 1)\delta\xi = 0 \qquad \forall \delta\xi$$

Then,

$$divu = \nabla \cdot u = f - 1 \tag{4.6.1}$$

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

$$= \int_\Omega (\nabla \times u - g)\delta\eta = 0 \qquad \forall \delta\eta$$

Then,

$$curlu = \nabla \times u = g \tag{4.6.2}$$

The state equations are given by:

$$divu = \nabla \cdot u = u_{1x_1} + u_{2x_2} = f - 1$$
$$curlu = \nabla \times u = u_{2x_1} - u_{1x_2} = g$$

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

$$L_{u_1} = \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} \left[\begin{array}{l} \left[\frac{1}{2} \int_\Omega T(x + (u_1 + \varepsilon\partial u_1, u_2) - R(x)) \right]^2 dx + \int_\Omega \xi(div(u_1 + \varepsilon\delta u_1, u_2) - f + 1) dx \\ + \int_\Omega \eta(curl(u_1 + \varepsilon\partial u_1, u_2) - g) dx \end{array} \right]$$

$$= \int_\Omega T\left(x + u(x) - R(x)\right) T_{\phi_1} \delta u_1 + \int_\Omega \xi \left(\delta u_1\right)_{x_1} + \int_\Omega \eta \left(-\delta u\right)_{x_2}$$

$$= \int_\Omega T\left(x + u(x) - R(x)\right) T_{\phi_1} \delta u_1 + \int_\Omega \left(\xi, -\eta\right) \cdot \nabla \delta u_1$$

$$= \int_\Omega \left[\left(T - R\right) T_{\phi_1} \delta u_1 - \nabla \cdot \left(\xi, -\eta\right) \delta u_1\right] \qquad \left(by \;\; Lemma \;\; 4.5.2 \;\; and \;\; h = \delta u_1 = 0 \;\; on \;\; \partial\Omega\right)$$

$$= \int_\Omega \left[\left(T - R\right) T_{\phi_1} - \nabla \cdot \left(\xi, -\eta\right)\right] \delta u_1 = 0 \qquad \forall \delta u_1$$

This gives us the first costate equation:

$$\nabla \cdot \left(\xi, -\eta\right) = \left(T - R\right) T_{\phi_1} \qquad (4.6.3)$$

$$L_{u_2} = \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} \left[\begin{array}{l} \left[\frac{1}{2} \int_\Omega T\left(x + \left(u_1, u_2 + \varepsilon \delta u_2\right) - R(x)\right)\right]^2 dx + \int_\Omega \xi \left(div\left(u_1, u_2 + \varepsilon \delta u_2\right) - f + 1\right) dx \\ + \int_\Omega \eta \left(curl\left(u_1, u_2 + \varepsilon \delta u_2\right) - g\right) dx \end{array}\right]$$

$$= \int_\Omega T\left(x + u(x) - R(x)\right) T_{\phi_2} \delta u_2 + \int_\Omega \xi \left(\delta u_2\right)_{x_2} + \int_\Omega \eta \left(\delta u_2\right)_{x_1}$$

$$= \int_\Omega T\left(x + u(x) - R(x)\right) T_{\phi_2} \delta u_2 + \int_\Omega \left(\eta, \xi\right) \cdot \nabla \delta u_2$$

$$= \int_\Omega \left[\left(T - R\right) T_{\phi_2} \delta u_2 - \nabla \cdot \left(\eta, \xi\right) \delta u_2\right] \qquad \left(by \;\; Lemma \;\; 4.5.2 \;\; and \;\; h = \delta u_2 = 0 \;\; on \;\; \partial\Omega\right)$$

$$= \int_\Omega \left[\left(T - R\right) T_{\phi_2} - \nabla \cdot \left(\eta, \xi\right)\right] \delta u_2 = 0 \qquad \forall \delta u_2$$

This gives us the second costate equation:

$$\nabla \cdot \left(\eta, \xi\right) = \left(T - R\right) T_{\phi_2} \qquad (4.6.4)$$

Optimality Conditions: The optimality conditions are obtained by solving the equations $L_f = 0$ and $L_g = 0$

$L_f = 0$

$$L_f = \frac{d}{d\varepsilon}\Big|_{\varepsilon=0}\left[\frac{d_1}{2}\int_\Omega (f+\varepsilon\delta f)^2 + \int_\Omega \xi\left(divu - (f+\varepsilon\delta f)+1\right)\right] = 0$$

$$L_f = \int_\Omega \left(d_1 f\delta f + \xi\delta f\right) = 0$$

$$L_f = \int_\Omega \left(d_1 f + \xi\right)\delta f = 0 \qquad \forall \delta f$$

This gives the first optimality condition:

$$d_1 f + \xi = 0 \tag{4.6.5}$$

$$L_g = 0$$

$$L_g = \frac{d}{d\varepsilon}\Big|_{\varepsilon=0}\left[\frac{d_2}{2}\int_\Omega (g+\varepsilon\delta g)^2 + \int_\Omega \eta\left(curlu - (g+\varepsilon\delta g)\right)\right] = 0$$

$$L_g = \int_\Omega \left(d_2 g - \eta\right)\delta g = 0 \qquad \forall \delta g$$

This gives the second optimality condition:

$$d_2 g - \eta = 0 \tag{4.6.6}$$

Now we write the optimality system as follows:

State Equations:

$$divu = \nabla \cdot u = f - 1$$
$$curlu = \nabla \times u = g$$

Costate Equations:

$$\nabla \cdot \left(\xi, -\eta\right) = \left(T - R\right)T_{\phi_1}$$

$$\nabla \cdot \left(\eta, \xi\right) = \left(T - R\right)T_{\phi_2}$$

Optimality conditions:

$$d_1 f = -\xi$$

$$d_2 g = \eta$$

Details of how to solve these numerically by multi-grid optimization were discussed in [11].

CHAPTER 5

CONCLUSIONS AND FUTURE WORK

In this dissertation we have shown that orthogonality of grids can be improved by a variational method. We do this without changing the cell size distribution of the adaptive grids.

A smoothness functional is added because the orthogonality problem is ill-posed.

We optimize a weighted sum of orthogonality functional and the smoothness functional. This procedure involves solving for a displacement u from the Euler-Lagrange equations.

The concern of grid folding is prevented by the grid deformation method and nearly orthogonal grids were obtained for λ between $0.6 \le \lambda \le 1.0$ and μ between $0 \le \mu \le 1$.

In the future we will find the maximum distortion angle between grids to quantify the results obtained. We will also try to determine if it is possible to find λ and μ theoretically.

In the future we will also include the div-curl system used in the image registration problem as our constraint.

We will minimize (5.1)

$$I = \frac{1}{2}\int_D \left(\varphi_{x_1} \cdot \varphi_{x_2}\right)^2 + \frac{\lambda}{2}\int_D \left|\nabla \varphi\right|^2 \qquad (5.1)$$

subject to:

$$\begin{aligned} div\,\varphi &= f - 1 \\ curl\,\varphi &= g \end{aligned} \qquad (5.2)$$

We will take f = 1, which means we keep the cell size distribution and only adjust the curl via the control variable g since the grid deformation method provides the control on size already.

As proven in Lemma 4.3.1, (5.2) is elliptic in the sense of Petrovski, so we will be able to omit the penalty term and minimize I in (5.1) subject to the div-curl constraint only.

Also in this dissertation, we have proven the existence of optimal solutions by the direct method in the calculus of variations for the image registration problem. We have discussed the non-uniqueness of optimal solutions since the image registration problem is ill-posed. We have proved the existence of Lagrange multipliers for the optimal solution. The Lagrange multipliers rule is established to turn the constrained optimization problem into an unconstrained optimization problem. We have derived the optimality system from which a numerical solution can be obtained. Overall, we combined the grid deformation method and Lagrange multiplier rule to solve the image registration problem.

In the future, we will prove that the optimal transformation Φ is close to $S^{-1}(x)$ if $\left|S^{-1}(x) - id\right| \le \varepsilon$ and $\left|\nabla S^{-1}(x)\right| \le \varepsilon$ .This is an important theoretical problem since S^{-1} is the ground truth (i.e. the exact solution) and we expect $\Phi \sim S^{-1}$.

REFERENCES

[1] G. Dela Pena "*Adaptive grid generation",* PhD thesis, UT Arlington, (1998).

[2] G. Liao and D. Anderson, *"A New Approach to grid Generation", Appl.Anal.* 44, (1992).

[3] G. Liao and J. Su, "*A direct method in Dacorogna-Moser's approach of grid generation problems", Appl. Anal.,* 45*,* (1993).

[4] G. Liao and J. Su, "*Grid Generation via Deformation",* Appl. Math. *Lett, 5, (1992).*

[5] G.Liao, T.Pan, and J. Su," *A Numerical Grid Generator based on Moser's Deformation method*", Numer.Meth.PDE, 10, (1994).

[6] 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).

[7] G.Liao and J. Su, *"A moving grid method for (1+1) dimension"*, Appl. Math Lett., 8, (1995).

[8] B.Semper and G.Liao*," A moving grid finite element method using grid deformation",* Numer.Meth.PDE, pp.11, (1995).

[9] B.Dacorogna and J Moser,"*On a PDE involving the Jacobian Determinant",* Ann.Inst H Poincare, 7, (1990).

[10] G. Liao, J.Xue, *Moving meshes by the Deformation Method, Special Issue: The International Symposium on Computing and Information, Journal of Computational and Applied Mathematics*, Volume 195, Issue 1, pp. 83-92, (2006).

[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] 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, pp* 26-38, (1993).

[17] Joe F. Thompson, Z.U.A. Warsi, C. Wayne Mastin, *Numerical Grid Generation*, Foundations and Applications, Elsevier Science Publishing Co., (1985).

[18] G. Liao, *Variational Approach to Grid Generation,* Numerical methods for Partial differential equations, 8,pp.143-147, (1992).

[19] J.U. Brackbill and J.S. Saltzman,"*Adaptive zoning for singular problems in two dimensions*,"J.Comput. Phys. 46, pp. 342-368, (1982).

[20] S. Steinberg and P. Roache," *Variational grid generation,"Numer.Meth.Partial Diff.Eq.2, pp* 71-96, (1985).

[21] J.Castillo, S.Steinberg, and P.J. Roache, "*Parameter estimation in variational grid generation*," Appl. Math. Comput. 28, pp. 1-23, (1988).

[22] R. Courant, Trans.Amer.Math.Soc. 50, pp.40, (1941).

[23] J.F. Thompson, F.C. Thames, and C.W. Masten, J.Comput.Phys. 15,pp. 299, (1974).

[24] 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).

[25] 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).

[26] P.A. Viola, *Alignment by Maximization of Mutual Information,* Ph.D thesis, Massachusetts Institute of Technology. (1995).

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

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

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

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

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

[32] J.Moser, *"Volume Elements of a Riemann Manifold",* Trans AMS, pp.120, (1965).

[33] E. Zeidler, *Nonlinear analysis and its Applications,* vol. III, Springer, New York, (1988).

BIOGRAPHICAL INFORMATION

Stephen Taiwo Salako was born in Westminster, England, United Kingdom. He earned his Bachelor's degree in Chemical Engineering with Minerals Engineering from The University of Birmingham, Birmingham, U.K. in July 2002 and then moved to Texas in 2003 to attend The University of Texas at San Antonio to earn a Master of Science degree in Mathematics in December 2005.

In 2006 he moved to Arlington, Texas and obtained a PhD degree in Mathematics from The University of Texas at Arlington in December 2009.

He has been a Graduate Teaching Assistant for 6 years at UT San Antonio and UT Arlington.

His research interests are Mathematics of Imaging, Optimization Problems, Partial Differential Equations, computational Mathematics, Chemical Plant Design, Minerals Processing and Solids Processing.