J Math Imaging Vis (2010) 36: 69–80
DOI 10.1007/s10851-009-0172-z

An Optimal Control Formulation of an Image Registration Problem

Eunjung Lee · Max Gunzburger

Published online: 8 October 2009

Abstract The basic idea of image registration is to find a reasonable transformation of an image so that the resulting difference between it and another image is made small. We derive an optimal control method for determining such a transformation; the approach is based on the grid deformation method and seeks to minimize an objective functional that measures the difference between the transformed image and the reference image. The existence of an optimal transformation is proved as is the applicability of the Lagrange multiplier method. Then, an optimality system from which optimal transformations can be obtained is derived.

Keywords Image registration · Optimal control problem · Grid deformation method · Optimal solution · Lagrange multipliers

1 Introduction

Image registration problems have been used in a wide variety of applications [2, 3]. In this paper, we consider image registration problems from different modalities. The images are generally obtained by different methods or equipments and represent different information of the same object or of different objects. Also, these images can be obtained by the same equipment but at different times. The aim of image registration is to find a reasonable geometric transformation such that the transformed template image becomes similar to the reference image.

Many fundamental approaches for image registration are introduced in [20]. A variational formulation of the image registration problem is widely used to find such a transformation in a number of approximate methods [7], for example, the Bayesian approach in [9] and curvature based algorithms in [10, 15]. Many of these approaches require high regularities in the control variables. For practicality, an octree method, a multilevel method, and many others were introduced to reduce the amount of data and thus the computational cost (see [12, 14, 16], and the references therein). As a different approach, image registration was considered in the context of a Mumford-Shah level-set method in [8].

Here, in order to find an optimal transformation that minimizes the difference between the transformed template image and the reference image, we adopt the grid deformation method [6, 17, 22]. The grid deformation method gives direct control over the cell size of the adaptive grid and determines the node velocities directly. It inherently defines a transformation which is obtained in two steps; in the first step, a given function is used to construct a vector field that satisfies a div-curl system and in the second step, this vector field is used to generate a transformation that moves the grid. When the transformation is diffeomorphic, the problem is called diffeomorphic registration [19, 20]. The grid deformation method introduced in [22], in fact, generates a diffeomorphic transformation. Now, control and optimization theories are required here to minimize a measure of the dissimilarity between the transformed version of image and the reference image by using the transformation obtained from

E. Lee (✉) · M. Gunzburger
Department of Scientific Computing, Florida State University, Dirac Science Library, Tallahassee, FL 32306-4120, USA
e-mail: elee4@fsu.edu

M. Gunzburger
e-mail: gunzburg@fsu.edu

Present address:
E. Lee
Department of Computational Science and Engineering, Yonsei University, Seoul, Republic of Korea
e-mail: eunjunglee@yonsei.ac.kr

the grid deformation method. There are various approaches to the solution of optimal control and optimization problems [11]. In this paper, we apply the Lagrange multiplier method to get the optimality system, the state and adjoint systems and optimality conditions (see [4, 13, 23]. Since the resulting system is constituted of several coupled systems, it is much efficient to solve the optimality system using optimization iteration than solve the system all at once.

In Sect. 1.1, we give a brief introduction of the grid deformation method. The optimal control formulation of the image registration problem is given in Sect. 1.2. In Sect. 2, we represent the optimal control problem induced from Sect. 1.2 and show the existence of the optimal solution to the optimal control problem. The existence of the Lagrange multiplier is given in Sect. 3 and we summarize the optimality system in Sect. 4.

1.1 The Grid Deformation Method

Let $\Omega \subset \mathbb{R}^2$ denote a bounded domain that is a convex polygon or that has a $\mathcal{C}^{1,1}$ boundary Γ. Let $\mathrm{R}(\mathbf{x})$ and $\mathrm{T}(\mathbf{x})$ for $\mathbf{x} \in \Omega$ denote two images; $\mathrm{R}(\mathbf{x})$ is kept fixed and is referred to as the *reference image* while $\mathrm{T}(\mathbf{x})$ is the image to be registered and is referred to as the *template image*. The image registration problem is to find an mapping $\tilde{\boldsymbol{\phi}}(\mathbf{x}) : \Omega \mapsto \Omega$ such that the transformed template image $\mathrm{T}(\tilde{\boldsymbol{\phi}}(\mathbf{x}))$ is "similar" to the reference image $\mathrm{R}(\mathbf{x})$. Here, we use an optimal control strategy based on the grid deformation method [6, 17, 22] to find this transformation.

Throughout, ∇ and $\nabla\cdot$ denote the spatial gradient and divergence operators, respectively; in two dimensions, we have the two curl operators $\nabla\times$ and $\nabla^{\perp}$, the first operating on 2-vectors and the second on scalars. Let $H^m(\Omega)$ denote the standard Sobolev space and $[H^m(\Omega)]^2$ denote the corresponding space of vector-valued functions each of whose components belong to $H^m(\Omega)$; for both scalar and vector-valued functions, $\|\cdot\|_m$ and $\langle\cdot,\cdot\rangle_m$ denote the corresponding norm and inner product, respectively. For $L^2(\Omega)$, we omit the subscript in the norm and inner product notation and $\langle\cdot,\cdot\rangle$ also denotes the duality pairing between spaces and their duals.

First, we briefly describe the grid deformation method; see [17] for details. Let $f(\mathbf{x})$ denote a positive function defined on Ω that satisfies $\int_\Omega (f(\mathbf{x})-1)d\mathbf{x} = 0$. Then, there exists a one-to-one transformation $\tilde{\boldsymbol{\phi}} : \Omega \mapsto \Omega$ such that

$$\det \nabla\tilde{\boldsymbol{\phi}}(\mathbf{x}) = f(\mathbf{x}) \quad \text{in } \Omega. \tag{1}$$

The grid deformation method generates, by using $f(\mathbf{x}) > 0$, a time-dependent mapping $\boldsymbol{\phi}(t,\mathbf{x})$ that moves the points in Ω in a desired way. Such a $\boldsymbol{\phi}(t,\mathbf{x})$ can be obtained by solving, for every $\mathbf{x} \in \Omega$, the following nonlinear ordinary differential equation

$$\begin{cases} \frac{\partial}{\partial t}\boldsymbol{\phi}(t,\mathbf{x}) = \mathbf{h}(t,\boldsymbol{\phi}(t,\mathbf{x})) & 0 < t \le 1, \\ \boldsymbol{\phi}(0,\mathbf{x}) = \mathbf{x}, & \\ \mathbf{h}(t,\mathbf{x}) = \frac{\mathbf{u}(\mathbf{x})}{t+(1-t)f(\mathbf{x})} & 0 < t \le 1, \end{cases} \tag{2}$$

where $\mathbf{u}$ satisfies

$$\begin{cases} \nabla\cdot\mathbf{u} = f-1 & \text{in } \Omega, \\ \nabla\times\mathbf{u} = \mathbf{0} & \text{in } \Omega, \\ \mathbf{n}\cdot\mathbf{u} = 0 & \text{on } \Gamma \end{cases} \tag{3}$$

with $\mathbf{n}$ the outward unit normal vector. Then, the transformed point $\tilde{\boldsymbol{\phi}}(\mathbf{x})$ in (1) is given by $\boldsymbol{\phi}(1,\mathbf{x})$ from (2), i.e., it can be shown that

$$\det \nabla\boldsymbol{\phi}(1,\mathbf{x}) = f(\mathbf{x}) \quad \text{in } \Omega, \tag{4}$$

i.e., by solving (2) and (3) for a given f, the grid deformation method provides a transformation $\boldsymbol{\phi}(t,\mathbf{x})$ satisfying (4). In adaptive mesh generation applications, the grid deformation method is used as follows. Suppose that we know some positive error estimator $e(\mathbf{x})$; we then define $f(\mathbf{x}) = c_1/e(\mathbf{x})$, where c_1 is a normalizing constant that ensures that $\int_\Omega (f(\mathbf{x})-1)d\mathbf{x} = 0$. Then, (4) implies that the grid deformation method produces a transformation such that the transformed grid has a small grid size if the error estimator is large and a large grid size if the error estimator is small.

1.2 Optimal Control Formulation of the Image Registration Problem

We now use the grid deformation method for the image registration problem. We simplify the grid deformation method by letting $\mathbf{h}(t,\mathbf{x}) = \mathbf{h}(1,\mathbf{x}) = \mathbf{u}(\mathbf{x})$ [18]. To achieve the goal of making the transformed template image close to the reference image, we seek a mapping $\boldsymbol{\phi}(t,\mathbf{x})$ that minimizes the L^2-norm of the difference between $\mathrm{T}(\boldsymbol{\phi}(1,\mathbf{x}))$ and $\mathrm{R}(\mathbf{x})$ over Ω. Specifically, we define the functional

$$\begin{aligned} J(\boldsymbol{\phi}|_{t=1}, f, g) = {} & \frac{1}{2}\|\mathrm{T}(\boldsymbol{\phi}(1,\mathbf{x})) - \mathrm{R}\|^2 + \frac{\alpha_{f_0}}{2}\|f\|^2 \\ & + \frac{\alpha_{f_1}}{2}\|\nabla f\|^2 + \frac{\alpha_{g_0}}{2}\|g\|^2 + \frac{\alpha_{g_1}}{2}\|\nabla g\|^2 \\ & - \beta\int_\Omega \log f\, d\mathbf{x}, \end{aligned} \tag{5}$$

where $\alpha_{f_0}, \alpha_{f_1}, \alpha_{g_0}$, and α_{g_1} are penalty parameters and β is a barrier parameter. We then seek controls $f(\mathbf{x})$ and $g(\mathbf{x})$ and states $\boldsymbol{\phi}(t,\mathbf{x})$ and $\mathbf{u}(\mathbf{x})$ such that $J(\cdot,\cdot,\cdot)$ is minimized,

subject to the constraints

$$\begin{cases} \nabla \cdot \mathbf{u}(\mathbf{x}) = f(\mathbf{x}) - 1 & \text{in } \Omega, \\ \nabla \times \mathbf{u}(\mathbf{x}) = g(\mathbf{x}) & \text{in } \Omega, \\ \frac{\partial \boldsymbol{\phi}(t,\mathbf{x})}{\partial t} = \mathbf{u}(\boldsymbol{\phi}(t,\mathbf{x})) & \text{in } (0,1] \times \Omega, \\ \int_\Omega (f(\mathbf{x}) - 1) d\mathbf{x} = 0, & \\ \mathbf{n} \cdot \mathbf{u}(\mathbf{x}) = 0 & \text{on } \Gamma, \\ \boldsymbol{\phi}(0,\mathbf{x}) = \mathbf{x} & \text{in } \Omega. \end{cases} \tag{6}$$

We also require that $\boldsymbol{\phi}(t,\mathbf{x}) \in (0,1] \times \Gamma$ for $\mathbf{x} \in \Gamma$.

The goal here is, of course, to minimize $\|\mathrm{T}(\boldsymbol{\phi}(1,\mathbf{x})) - \mathrm{R}\|$; however, minimizing only this term can result in unbounded optimal control functions because of the lack of any explicit dependence of the objective functional on the controls. Therefore, we penalize the objective functional with H^1-norms of f and g. With such penalization, we are able to show the existence of optimal solutions. If, instead, we use weaker L^2-norm penalization, we have not been able to prove this result; furthermore, computational studies indicate that L^2-norm penalization may well not be sufficient to guarantee the existence of optimal solutions.

We also need $f(\mathbf{x}) > 0$ in Ω. To enforce this constraint, we introduced, in (5), the integral of the logarithmic barrier function $-\log f$. If the H^1-norm of f is bounded, then, in two dimensions, $|f|$ is bounded above *a.e.* Thus, limiting $-\beta \int_\Omega \log f d\mathbf{x}$ to be bounded yields that f is bounded below by 0 *a.e.* in Ω.

2 The Optimal Control Problem and the Existence of Optimal Solutions

We define the product Hilbert spaces

$$V = [H^2(\Omega)]^2 \times [H^1((0,1); L^2(\Omega))]^2 \times H^1(\Omega) \times H^1(\Omega),$$

$$W = H^1(\Omega) \times H^1(\Omega) \times [L^2((0,1) \times \Omega)]^2 \times \mathbb{R} \times H^{\frac{3}{2}}(\Gamma) \times L^2(\Omega)$$

and define the operator $M : V \mapsto W$ as follows:

$$M(\mathbf{u}, \boldsymbol{\phi}, f, g) = \mathbf{0} \tag{7}$$

if and only if

$$\langle \nabla \cdot \mathbf{u} - f + 1, \xi \rangle = 0 \quad \forall \xi \in H^{-1}(\Omega), \tag{8}$$

$$\langle \nabla \times \mathbf{u} - g, \eta \rangle = 0 \quad \forall \eta \in H^{-1}(\Omega), \tag{9}$$

$$\left\langle \frac{\partial \boldsymbol{\phi}}{\partial t} - \mathbf{u}(\boldsymbol{\phi}), \boldsymbol{\psi} \right\rangle = 0 \quad \forall \boldsymbol{\psi} \in [L^2((0,1) \times (\Omega))]^2, \tag{10}$$

$$\sigma \int_\Omega (f - 1) d\mathbf{x} = 0 \quad \forall \sigma \in \mathbb{R}, \tag{11}$$

$$\langle \mathbf{n} \cdot \mathbf{u}, \nu \rangle = 0 \quad \forall \nu \in L^2((0,1); H^{-\frac{3}{2}}(\Gamma)), \tag{12}$$

$$\langle \boldsymbol{\phi}(0,\mathbf{x}) - \mathbf{x}, \boldsymbol{\mu} \rangle = 0 \quad \forall \boldsymbol{\mu} \in [L^2(\Omega)]^2. \tag{13}$$

We recognize (8)–(13) as a weak formulation of the constraint equations (6). If we assume that $\mathrm{T}(\mathbf{x})$ and $\mathrm{R}(\mathbf{x})$ belong to $[H^1(\Omega)]^2$, then the set of all admissible solutions is defined by

$$A_{ad} = \{(\mathbf{u}, \boldsymbol{\phi}, f, g) \in V \mid J(\boldsymbol{\phi}|_{t=1}, f, g) \text{ is bounded and (7) is satisfied}\}. \tag{14}$$

Then, the optimal control problem is given by:

$$\begin{cases} \text{find } (\mathbf{u}, \boldsymbol{\phi}, f, g) \in A_{ad} \\ \text{which minimizes the functional (5).} \end{cases} \tag{15}$$

2.1 Existence of Optimal Solutions

We next show the existence of solutions of the optimal control problem (15).

Under the assumptions we have made about $\Omega \subset \mathbb{R}^2$, it is well-known that, given $f, g \in H^1(\Omega)$, there exists a $\mathbf{u} \in [H^2(\Omega)]^2$ satisfying (8), (9), and (12), i.e.,

$$\begin{aligned} &\nabla \cdot \mathbf{u} = f - 1 \quad \text{and} \quad \nabla \times \mathbf{u} = g \quad \text{in } \Omega, \\ &\mathbf{n} \cdot \mathbf{u} = 0 \quad \text{on } \Gamma. \end{aligned} \tag{16}$$

Since $\mathbf{u} \in [H^2(\Omega)]^2$, by the Sobolev embedding theory, $\mathbf{u}$ is bounded and Hölder-continuous, that is, for all $\mathbf{x}, \mathbf{y} \in \Omega$,

$$\text{for some constant } K_1 > 0, \quad |\mathbf{u}(\mathbf{x})| \le K_1 \tag{17}$$

and

$$\begin{cases} \text{for some constant } K_2 > 0, \\ |\mathbf{u}(\mathbf{x}) - \mathbf{u}(\mathbf{y})| \le K_2 |\mathbf{x} - \mathbf{y}|^\lambda, \quad \text{for } 0 < \lambda < 1. \end{cases} \tag{18}$$

Given such a $\mathbf{u} \in [H^2(\Omega)]^2$ and $\boldsymbol{\phi}_0(\mathbf{x}) \in [L^2(\Omega)]^2$, consider the nonlinear ordinary differential equation

$$\begin{aligned} &\frac{\partial \boldsymbol{\phi}(t,\mathbf{x})}{\partial t} = \mathbf{u}(\boldsymbol{\phi}(t,\mathbf{x})) \quad \text{in } (0,1] \times \Omega, \\ &\boldsymbol{\phi}(0,\mathbf{x}) = \boldsymbol{\phi}_0(\mathbf{x}) \quad \text{in } \Omega. \end{aligned} \tag{19}$$

Since $\mathbf{n} \cdot \mathbf{u} = 0$ on Γ, we have that

$$\frac{\partial}{\partial t}(\mathbf{n} \cdot \boldsymbol{\phi}) = 0 \quad \text{in } (0,1] \times \Gamma$$

so that, if $\boldsymbol{\phi}(0,\mathbf{x}) \in \Omega$, we have that

$$\boldsymbol{\phi}(t,\mathbf{x}) \in \Omega \quad \text{for all } t \ge 0. \tag{20}$$

We then have

$$|\boldsymbol{\phi}(t,\mathbf{x}) - \boldsymbol{\phi}_0(\mathbf{x})| \le K_3 = |\Omega| \quad \text{for all } t \ge 0, \tag{21}$$

where $|\Omega|$ denotes the diameter of Ω. Note also that if $\boldsymbol{\phi}(0,\mathbf{x}) \in \Gamma$ for some $\mathbf{x}$, then $\boldsymbol{\phi}(t,\mathbf{x}) \in \Gamma$ as well for $t > 0$.

The existence of a solution of the system (19) can be shown using the Peano existence theorem. Here, we modify for our setting the proof of that theorem given in [5]. To this end, we first quote two well-known results.

Lemma 1 [24] *Let X, Y denote two Banach spaces. If the embedding $X \subseteq Y$ is compact, then, as $n \to \infty$,*

$$u_n \to u \text{ weakly in } X \quad \text{implies} \quad u_n \to u \text{ strongly in } Y.$$

Lemma 2 [1] *Let Ω be open an bounded Lipschitz continuous subset of $\mathbb{R}^d$. Then, $L^p(\Omega)$ is compactly embedded into $L^{p'}(\Omega)$ if $1 \le p' < p$.*

Proposition 1 *Let $\mathbf{u}$ satisfy (16) so that (17), (18), and (20) hold. Then, (19) has at least one solution belonging to $[L^2((0,1)\times\Omega)]^2$.*

Proof Since $\mathcal{C}_0^\infty((0,1)\times\Omega)$ is dense in $L^2((0,1)\times\Omega)$, we choose the test vector functions $\boldsymbol{\psi}$ in (10) from $[\mathcal{C}_0^\infty((0,1)\times\Omega)]^2$, i.e., we solve

$$\left\langle \frac{\partial \boldsymbol{\phi}}{\partial t}, \boldsymbol{\psi} \right\rangle = \langle \mathbf{u}(\boldsymbol{\phi}), \boldsymbol{\psi} \rangle \quad \forall \boldsymbol{\psi} \in [\mathcal{C}_0^\infty((0,1)\times\Omega)]^2 \tag{22}$$

for given $\mathbf{u}$. Equation (22) has the equivalent integral equation form

$$\langle \boldsymbol{\phi}(t,\cdot), \boldsymbol{\psi} \rangle = \left\langle \boldsymbol{\phi}_0(\mathbf{x}) + \int_0^t \mathbf{u}(\boldsymbol{\phi}(s,\mathbf{x}))ds, \boldsymbol{\psi} \right\rangle \tag{23}$$

for all $\boldsymbol{\psi} \in [\mathcal{C}_0^\infty((0,1)\times\Omega)]^2$. Let $T_1 = \min\{1, K_3/K_1\}$ and assume that $0 \le t \le T_1$. For any $\mathbf{x} \in \Omega$ and $m = 1, 2, \ldots$, we construct a sequence of functions $\boldsymbol{\phi}^m(t,\mathbf{x})$ as follows:

$$\boldsymbol{\phi}^m(t,\mathbf{x}) = \begin{cases} \boldsymbol{\phi}_0(\mathbf{x}) & \text{for } 0 \le t \le \frac{T_1}{m}, \\ \boldsymbol{\phi}_0(\mathbf{x}) + \int_0^{t-\frac{T_1}{m}} \mathbf{u}(\boldsymbol{\phi}^m(s,\mathbf{x}))ds & \\ \quad \text{for } \frac{T_1}{m} < t \le T_1. \end{cases} \tag{24}$$

This formula defines the values of $\boldsymbol{\phi}^m(t,\mathbf{x})$ for $t \in (T_1/m, T_1]$ in terms of the previous values of $\boldsymbol{\phi}^m(s,\mathbf{x})$ for $0 \le s \le t - T_1/m$. Then, it follows from (21) (see [5, p. 158]) that $\boldsymbol{\phi}^m(t,\mathbf{x})$ is defined for $0 \le t \le T_1$. Since we have

$$\begin{aligned} |\boldsymbol{\phi}^m(t,\mathbf{x})| &\le |\boldsymbol{\phi}_0(\mathbf{x})| + \int_0^{t-\frac{T_1}{m}} |\mathbf{u}(\boldsymbol{\phi}^m(s,\mathbf{x}))|ds \\ &\le |\boldsymbol{\phi}_0(\mathbf{x})| + T_1 K_1 \quad \text{for } t \in (0, T_1), \end{aligned}$$

the $L^2((0,T_1)\times\Omega)$-norm of $\boldsymbol{\phi}^m$ is uniformly bounded so that there exists a weakly convergent subsequence $\{\boldsymbol{\phi}^{m_i}\}$ such that

$$\boldsymbol{\phi}^{m_i} \longrightarrow \boldsymbol{\phi}^\infty \quad \text{weakly in } [L^2((0,T_1)\times\Omega)]^2 \tag{25}$$

as $m_i \to \infty$.

We next need to verify that the limit function $\boldsymbol{\phi}^\infty$ satisfies (23) for $t \le T_1$. For $t \in ((T_1/m), T_1]$, (24) can be written as

$$\begin{aligned} \boldsymbol{\phi}^m(t,\mathbf{x}) &= \boldsymbol{\phi}_0(\mathbf{x}) + \int_0^t \mathbf{u}(\boldsymbol{\phi}^m(s,\mathbf{x}))ds \\ &\quad - \int_{t-\frac{T_1}{m}}^t \mathbf{u}(\boldsymbol{\phi}^m(s,\mathbf{x}))ds. \end{aligned}$$

Hence, we have, for $\boldsymbol{\psi} \in [\mathcal{C}_0^\infty((0,1)\times\Omega)]^2$,

$$\begin{aligned} &\langle \boldsymbol{\phi}^{m_i}, \boldsymbol{\psi} \rangle \\ &= \int_\Omega \left(\int_0^{\frac{T_1}{m_i}} \boldsymbol{\phi}_0(\mathbf{x}) \cdot \boldsymbol{\psi} dt \right. \\ &\quad \left. + \int_{\frac{T_1}{m_i}}^{T_1} \left(\boldsymbol{\phi}_0(\mathbf{x}) + \int_0^{t-\frac{T_1}{m_i}} \mathbf{u}(\boldsymbol{\phi}^{m_i}(s,\mathbf{x}))ds \right) \cdot \boldsymbol{\psi} dt \right) d\mathbf{x} \\ &= \int_\Omega \int_0^{T_1} \left(\boldsymbol{\phi}_0(\mathbf{x}) + \int_0^t \mathbf{u}(\boldsymbol{\phi}^{m_i}(s,\mathbf{x}))ds \right) \cdot \boldsymbol{\psi} dt d\mathbf{x} \\ &\quad - \int_\Omega \left(\int_0^{\frac{T_1}{m_i}} \int_0^t \mathbf{u}(\boldsymbol{\phi}^{m_i}(s,\mathbf{x}))ds \cdot \boldsymbol{\psi} dt \right. \\ &\quad \left. + \int_{\frac{T_1}{m_i}}^{T_1} \int_{t-\frac{T_1}{m_i}}^t \mathbf{u}(\boldsymbol{\phi}^{m_i}(s,\mathbf{x}))ds \cdot \boldsymbol{\psi} dt \right) d\mathbf{x}. \end{aligned} \tag{26}$$

Owing to $\boldsymbol{\psi} \in [\mathcal{C}_0^\infty((0,1)\times\Omega)]^2$, $\boldsymbol{\psi}$ satisfies $|\boldsymbol{\psi}| \le K_4$ for some constant $K_4 > 0$. Then, for the second term in (26), we have that

$$\begin{aligned} &\int_\Omega \int_0^{\frac{T_1}{m_i}} \int_0^t \mathbf{u}(s, \boldsymbol{\phi}^{m_i}(s,\mathbf{x}))ds \cdot \boldsymbol{\psi} dt d\mathbf{x} \\ &\quad \le \int_\Omega \int_0^{\frac{T_1}{m_i}} t K_1 |\boldsymbol{\psi}| dt d\mathbf{x} \le \frac{C}{m_i^2} \longrightarrow 0 \end{aligned}$$

as $m_i \to \infty$, where C only depends on K_1, K_3, K_4, and T_1. Similarly, the third term in (26) is less than $(K_1 K_3 K_4 T_1)/m_i$ so that it goes to zero as $m_i \to \infty$. Thus, in order to verify that $\boldsymbol{\phi}^\infty$ satisfies (23), it suffices to show that

$$\begin{aligned} &\left\langle \boldsymbol{\phi}_0(\mathbf{x}) + \int_0^t \mathbf{u}(\boldsymbol{\phi}^{m_i}(s,\mathbf{x}))ds, \boldsymbol{\psi} \right\rangle \\ &\quad \longrightarrow \left\langle \boldsymbol{\phi}_0(\mathbf{x}) + \int_0^t \mathbf{u}(\boldsymbol{\phi}^\infty(s,\mathbf{x}))ds, \boldsymbol{\psi} \right\rangle \end{aligned} \tag{27}$$

as $m_i \to \infty$.

From (25) and Lemmas 1 and 2, we have that $\boldsymbol{\phi}^{m_i}$ converges to $\boldsymbol{\phi}^\infty$ strongly in $[L^{p'}((0,T_1)\times\Omega)]^2$ for $1 \le p' < 2$. Then, (18), the Cauchy inequality, and the above strong convergence yield

$$\left\langle \int_0^t \mathbf{u}(\boldsymbol{\phi}^{m_i}(s,\mathbf{x}))ds, \boldsymbol{\psi} \right\rangle - \left\langle \int_0^t \mathbf{u}(\boldsymbol{\phi}^\infty(s,\mathbf{x}))ds, \boldsymbol{\psi} \right\rangle$$

$$
\begin{aligned}
&= \int_\Omega \int_0^{T_1} \int_0^t \big(\mathbf{u}(\boldsymbol{\phi}^{m_i}(s,\mathbf{x})) - \mathbf{u}(\boldsymbol{\phi}^\infty(s,\mathbf{x}))\big) \\
&\quad \times \boldsymbol{\psi}(t,\mathbf{x}) ds dt d\mathbf{x} \\
&\le \int_\Omega \int_0^{T_1} \int_0^t K_2 |\boldsymbol{\phi}^{m_i}(s,\mathbf{x}) - \boldsymbol{\phi}^\infty(s,\mathbf{x})|^\lambda K_4 ds dt d\mathbf{x} \\
&\le C \left(\int_\Omega \int_0^{T_1} |\boldsymbol{\phi}^{m_i}(s,\mathbf{x}) - \boldsymbol{\phi}^\infty(s,\mathbf{x})|^{\lambda\cdot\frac{1}{\lambda}} ds d\mathbf{x} \right)^\lambda \longrightarrow 0
\end{aligned}
$$

as $m_i \to \infty$ where $0 < \lambda < 1$.

We follow [21] to complete the proof. If $T_1 < 1$, then we restart the ordinary differential equation in (19) at $t = T_1$ with initial condition given by $\boldsymbol{\phi}(T_1, \mathbf{x})$ to show the existence of a solution for $t \le T_2$ for some $T_2 > T_1$. Since (17), (18), and (20) hold for all $t > 0$, then after each restart t is advanced by a finite amount so that after a sufficient number of restarts, we will have existence for $t \in [0, 1]$. The proof is complete. □

Corollary 1 *Let the hypotheses of Proposition 1 hold and let $\phi_0(\mathbf{x}) \in H^1(\Omega)$. Then the solution $\boldsymbol{\phi}(t,\mathbf{x})$ of (19) also belongs to $[L^2((0,1); H^1(\Omega))]^2$.*

Proof We only sketch the proof. The technique used for this proof is explained in detail in the proof of Theorem 1. Take the gradient of both sides of (10) to obtain

$$
\frac{\partial \nabla \boldsymbol{\phi}}{\partial t} = \nabla \frac{\partial \boldsymbol{\phi}}{\partial t} = \nabla \boldsymbol{\phi} \nabla_{\boldsymbol{\phi}} \mathbf{u}(\boldsymbol{\phi}).
$$

Multiplying by $\nabla\boldsymbol{\phi}$ and integrating over $(0, 1)$ and Ω give

$$
\begin{aligned}
&\int_\Omega |\nabla \boldsymbol{\phi}|^2 d\mathbf{x} - \int_\Omega |\nabla \phi_0(\mathbf{x})|^2 d\mathbf{x} \\
&\quad = 2 \int_\Omega \int_0^1 \nabla \boldsymbol{\phi} \cdot \frac{\partial \nabla \boldsymbol{\phi}}{\partial s} ds d\mathbf{x} \\
&\quad \le 2 \int_0^1 \int_\Omega |\nabla \boldsymbol{\phi}|^2 |\nabla_{\boldsymbol{\phi}} \mathbf{u}(s,\boldsymbol{\phi})| d\mathbf{x} ds \\
&\quad \le c \int_0^1 \int_\Omega |\nabla \boldsymbol{\phi}|^2 d\mathbf{x} ds.
\end{aligned}
$$

By Gronwall's lemma, $\boldsymbol{\phi}(t,\cdot) \in [H^1(\Omega)]^2$ for any $t \in [0, 1]$. □

The following theorem shows the existence of optimal solutions.

Theorem 1 *There exists a solution $(\mathbf{u}, \boldsymbol{\phi}, f, g) \in A_{ad}$ for the optimal control problem (15).*

Proof We first choose $f = 1$ and $g = 0$. Then, we immediately see that $\mathbf{u} = \mathbf{0}$ and $\boldsymbol{\phi} = \mathbf{x}$ so that A_{ad} is not empty.

Let $\{(\mathbf{u}^n, \boldsymbol{\phi}^n, f^n, g^n)\} \subset A_{ad}$ denote a minimizing sequence, i.e., we have that

$$
\lim_{n\to\infty} J(\boldsymbol{\phi}^n|_{t=1}, f^n, g^n) = \inf_{(\mathbf{u},\boldsymbol{\phi},f,g)\in A_{ad}} J(\boldsymbol{\phi}|_{t=1}, f, g). \quad (28)
$$

By (14), f^n and g^n are bounded in $H^1(\Omega)$. Then, well-known regularity results yields $\mathbf{u}^n \in [H^2(\Omega)]^2$ and

$$
\|\mathbf{u}^n\|_2 \le c(\|f^n\|_1 + \|g^n\|_1).
$$

Since $\mathbf{u}^n \in [H^2(\Omega)]^2$, $\mathbf{u}^n$ is bounded, i.e., $|\mathbf{u}^n| \le K$ for some constant $K > 0$. The boundedness of $\mathbf{u}^n$ implies

$$
\int_\Omega \int_0^1 \left| \frac{\partial \boldsymbol{\phi}^n(t,\mathbf{x})}{\partial t} \right|^2 dt d\mathbf{x} = \int_\Omega \int_0^1 |\mathbf{u}^n(\boldsymbol{\phi}^n(t,\mathbf{x}))|^2 dt d\mathbf{x} \le C
$$

and, therefore, $\partial\boldsymbol{\phi}^n/\partial t \in [L^2((0,1);\Omega)]^2$. Integrating the equality

$$
\frac{d}{ds} \int_\Omega (\boldsymbol{\phi}^n(s,\mathbf{x}))^2 d\mathbf{x} = 2 \int_\Omega \boldsymbol{\phi}^n(s,\mathbf{x}) \frac{\partial \boldsymbol{\phi}^n(s,\mathbf{x})}{\partial s} d\mathbf{x}
$$

over $(0, t)$ implies

$$
\begin{aligned}
&\int_0^t \frac{d}{ds} \int_\Omega (\boldsymbol{\phi}^n(s,\mathbf{x}))^2 d\mathbf{x} ds \\
&\quad = 2 \int_0^t \int_\Omega \boldsymbol{\phi}^n(s,\mathbf{x}) \frac{\partial \boldsymbol{\phi}^n(s,\mathbf{x})}{\partial s} d\mathbf{x} ds \\
&\quad = 2 \int_0^t \int_\Omega \boldsymbol{\phi}^n(s,\mathbf{x}) \mathbf{u}^n(\boldsymbol{\phi}^n(s,\mathbf{x})) d\mathbf{x} ds \\
&\quad \le K_1 \int_0^t \int_\Omega (\boldsymbol{\phi}^n(s,\mathbf{x}))^2 d\mathbf{x} ds, \qquad (29)
\end{aligned}
$$

where K_1 is a positive constant independent of $\boldsymbol{\phi}^n$ and t. The last inequality results from the Hölder inequality and the boundedness of $\mathbf{u}^n$. Since $\boldsymbol{\phi}^n(0,\mathbf{x}) = \mathbf{x}$, the left-hand side of the (29) becomes

$$
\begin{aligned}
&\int_\Omega (\boldsymbol{\phi}^n(t,\mathbf{x}))^2 d\mathbf{x} - \int_\Omega (\boldsymbol{\phi}^n(0,\mathbf{x}))^2 d\mathbf{x} \\
&\quad = \int_\Omega (\boldsymbol{\phi}^n(t,\mathbf{x}))^2 d\mathbf{x} - K_2, \qquad (30)
\end{aligned}
$$

where the positive constant K_2 only depends on Ω. From (29) and (30), Gronwall's lemma yields

$$
\int_\Omega (\boldsymbol{\phi}^n(t,\mathbf{x}))^2 dx \le K_2 e^{K_1 t} \quad \forall t \in [0, 1].
$$

Therefore, we obtain that $\|\boldsymbol{\phi}^n\|_{L^2((0,1)\times\Omega)}$ is bounded. Then, we can choose subsequences such that

$$
\begin{aligned}
&\mathbf{u}^n \to \hat{\mathbf{u}} \quad \text{weakly in } [H^2(\Omega)]^2, \\
&\boldsymbol{\phi}^n \to \hat{\boldsymbol{\phi}} \quad \text{weakly in } [H^1((0,1); L^2(\Omega))]^2, \\
&f^n \to \hat{f} \quad \text{weakly in } H^1(\Omega), \\
&g^n \to \hat{g} \quad \text{weakly in } H^1(\Omega)
\end{aligned} \qquad (31)
$$

for some $(\hat{\mathbf{u}}, \hat{\boldsymbol{\phi}}, \hat{f}, \hat{g}) \in V$.

It can be easily shown that (31) implies that the limit $(\hat{\mathbf{u}}, \hat{f}, \hat{g})$ satisfies the constraint equations (8), (9), (11), (12), and (13). To prove that $\hat{\boldsymbol{\phi}}$ satisfies the third constraint equation (10) for a given $\hat{\mathbf{u}}$, we first choose the test function $\boldsymbol{\psi}$ from $[\mathcal{C}_0^\infty((0,1);\Omega)]^2$. Then, we show

$$\left\langle \frac{\partial \hat{\boldsymbol{\phi}}}{\partial t}, \boldsymbol{\psi} \right\rangle = \left\langle \hat{\mathbf{u}}(\hat{\boldsymbol{\phi}}), \boldsymbol{\psi} \right\rangle \quad \forall \boldsymbol{\psi} \in [\mathcal{C}_0^\infty((0,1);\Omega)]^2.$$

By (31), we have

$$\begin{aligned}
&\int_\Omega \int_0^1 \left(\frac{\partial \boldsymbol{\phi}^n}{\partial t} - \mathbf{u}^n(\boldsymbol{\phi}^n) - \left(\frac{\partial \hat{\boldsymbol{\phi}}}{\partial t} - \hat{\mathbf{u}}(\hat{\boldsymbol{\phi}}) \right) \right) \cdot \boldsymbol{\psi}\, dt d\mathbf{x} \\
&= \int_\Omega \int_0^1 \left(\frac{\partial \boldsymbol{\phi}^n}{\partial t} - \frac{\partial \hat{\boldsymbol{\phi}}}{\partial t} \right) \cdot \boldsymbol{\psi} \\
&\quad - \left(\mathbf{u}^n(\boldsymbol{\phi}^n) - \mathbf{u}^n(\hat{\boldsymbol{\phi}}) + \mathbf{u}^n(\hat{\boldsymbol{\phi}}) - \hat{\mathbf{u}}(\hat{\boldsymbol{\phi}}) \right) \cdot \boldsymbol{\psi}\, dt d\mathbf{x} \\
&\longrightarrow - \int_\Omega \int_0^1 (\mathbf{u}^n(\boldsymbol{\phi}^n) - \mathbf{u}^n(\hat{\boldsymbol{\phi}})) \cdot \boldsymbol{\psi}\, dt d\mathbf{x}. \qquad (32)
\end{aligned}$$

The Hölder continuity of $\mathbf{u}^n$ (see (18)), $|\boldsymbol{\psi}| \le K'$, and the Cauchy inequality yield

$$\begin{aligned}
&\int_\Omega \int_0^1 (\mathbf{u}^n(\boldsymbol{\phi}^n) - \mathbf{u}^n(\hat{\boldsymbol{\phi}})) \cdot \boldsymbol{\psi}\, dt d\mathbf{x} \\
&\le \int_\Omega \int_0^1 L|\boldsymbol{\phi}^n(t,\mathbf{x}) - \hat{\boldsymbol{\phi}}(t,\mathbf{x})|^\lambda K' dt d\mathbf{x} \\
&\le c \left(\int_\Omega \int_0^1 |\boldsymbol{\phi}^n(t,\mathbf{x}) - \hat{\boldsymbol{\phi}}(t,\mathbf{x})|^{\lambda \cdot \frac{1}{\lambda}} dt d\mathbf{x} \right)^\lambda \qquad (33)
\end{aligned}$$

for $0 < \lambda < 1$. Since $\boldsymbol{\phi}^n$ weakly converges to $\hat{\boldsymbol{\phi}}$ in $[H^1((0,1);L^2(\Omega))]^2$, Lemmas 1 and 2 imply $\boldsymbol{\phi}^n$ converges strongly to $\hat{\boldsymbol{\phi}}$ in $[L^1((0,1)\times\Omega)]^2$. Hence, the right-hand side of (33) goes to 0 and, therefore (32) goes to 0. Thus, $(\hat{\mathbf{u}}, \hat{\boldsymbol{\phi}}, \hat{f}, \hat{g}) \in V$ satisfies (7). By the weakly lower semicontinuity of J, we finally obtain

$$J(\hat{\boldsymbol{\phi}}|_{t=1}, \hat{f}, \hat{g}) = \inf_{(\mathbf{u},\boldsymbol{\phi},f,g)\in A_{ad}} J(\boldsymbol{\phi}|_{t=1}, f, g). \qquad \square$$

3 Existence of Lagrange Multipliers

In the previous section, we showed the existence of the optimal solutions of the optimal control problem (15). Now, we want to use the Lagrange multiplier rule to turn the constrained minimization problem (15) into an unconstrained problem and then to derive an optimality system. Thus, we first show the existence of proper Lagrange multipliers.

Let M be the constraint equations from V to W defined as in (7). Let $M'(\mathbf{u}, \boldsymbol{\phi}, f, g) \in \mathcal{L}(V; W)$ be the first Fréchet derivative of M at $(\mathbf{u}, \boldsymbol{\phi}, f, g)$ defined as

$$M'(\mathbf{u}, \boldsymbol{\phi}, f, g) \cdot (\tilde{\mathbf{u}}, \tilde{\boldsymbol{\phi}}, \tilde{f}, \tilde{g}) = \mathbf{F} \qquad (34)$$

for $(\tilde{\mathbf{u}}, \tilde{\boldsymbol{\phi}}, \tilde{f}, \tilde{g}) \in V$ and $\mathbf{F} \equiv (f_1, f_2, \mathbf{f}_3, f_4, f_5, \mathbf{f}_6) \in W$ if and only if

$$\begin{cases}
\langle \nabla \cdot \tilde{\mathbf{u}} - \tilde{f}, \xi \rangle = \langle f_1, \xi \rangle \quad \forall \xi \in H^{-1}(\Omega), \\
\langle \nabla \times \tilde{\mathbf{u}} - \tilde{g}, \eta \rangle = \langle f_2, \eta \rangle \quad \forall \eta \in H^{-1}(\Omega), \\
\langle \frac{\partial \tilde{\boldsymbol{\phi}}}{\partial t} - \tilde{\boldsymbol{\phi}} \cdot \nabla_{\boldsymbol{\phi}} \mathbf{u}(\boldsymbol{\phi}) - \tilde{\mathbf{u}}(\boldsymbol{\phi}), \boldsymbol{\psi} \rangle = \langle \mathbf{f}_3, \boldsymbol{\psi} \rangle \\
\quad \forall \boldsymbol{\psi} \in [L^2((0,1)\times\Omega)]^2, \\
\sigma \int_\Omega \tilde{f}\, d\mathbf{x} = \sigma f_4 \qquad \forall \sigma \in \mathbb{R}, \\
\langle \mathbf{n} \cdot \tilde{\mathbf{u}}, \nu \rangle = \langle f_5, \nu \rangle \quad \forall \nu \in H^{-\frac{3}{2}}(\Gamma), \\
\langle \tilde{\boldsymbol{\phi}}(0,\mathbf{x}), \boldsymbol{\mu} \rangle = \langle \mathbf{f}_6, \boldsymbol{\mu} \rangle \quad \forall \boldsymbol{\mu} \in [L^2(\Omega)]^2.
\end{cases} \qquad (35)$$

The following two lemmas are useful in proving the existence of Lagrange multipliers. Denote the Sobolev space $W_\infty^m(\Omega) = \{u : D^\alpha u \in L^\infty(\Omega), 0 \le |\alpha| \le m\}$.

Lemma 3 *If $u \in H^m(\Omega)$ and $\|u\|_m \le K_u$ for some constant $K_u > 0$, then $u \in W_\infty^m(\Omega)$, that is, there exists a positive constant L_u such that*

$$|D^\alpha u(\mathbf{x})| \le L_u, \quad 0 \le |\alpha| \le m \text{ a.e., in } \Omega. \qquad (36)$$

Proof Suppose $u \notin W_\infty^m(\Omega)$, then there exists a set $U \subset \Omega$ such that the measure of U, $m(U)$, is not 0 and $|D^\beta u(\mathbf{x})| \ge n$ for some $0 \le |\beta| \le m$ and for all $\mathbf{x} \in U$. Then

$$\begin{aligned}
&\int_\Omega \sum_{0\le|\alpha|\le m} |D^\alpha u(\mathbf{x})|^2 d\mathbf{x} \\
&\ge \int_U \sum_{0\le|\alpha|\le m} |D^\alpha u(\mathbf{x})|^2 d\mathbf{x} \\
&\ge n^2 \cdot m(U) \longrightarrow \infty, \quad \text{as } n \to \infty
\end{aligned}$$

which contradicts the fact that $\|u\|_m$ is bounded. $\square$

To show the existence of Lagrange multiplier, we first prove that the operator $M'(\mathbf{u}, \boldsymbol{\phi}, f, g)$ is onto a certain space. In the following lemma, we prove M' in a strong formulation is onto.

Lemma 4 *Let $(\mathbf{u}, \boldsymbol{\phi}, f, g)$ be a solution to the optimal control problem (15). Then, the operator $M'(\mathbf{u}, \boldsymbol{\phi}, f, g)$ is onto W.*

Proof Let $(f_1, f_2, \mathbf{f}_3, f_4, f_5, \mathbf{f}_6) \in W$. Since the system

$$M'(\mathbf{u}, \boldsymbol{\phi}, f, g) \cdot (\tilde{\mathbf{u}}, \tilde{\boldsymbol{\phi}}, \tilde{f}, \tilde{g}) = (f_1, f_2, \mathbf{f}_3, f_4, f_5, \mathbf{f}_6)$$

is underdetermined, we can choose some variables and then solve for the rest. We choose $\tilde{f} \in H^1(\Omega)$ such that

$$\int_\Omega \tilde{f} d\mathbf{x} = f_4$$

and choose any $\tilde{g} \in H^1(\Omega)$. For $f_5 \in H^{\frac{3}{2}}(\Gamma)$, by the trace theorem, there exists a $\mathbf{v} \in H^2(\Omega)$ such that $\mathbf{n} \cdot \mathbf{v} = f_5$. Then, we solve for $\mathbf{w}$ satisfying

$$\nabla \cdot \mathbf{w} = \tilde{f} + f_1 - \nabla \cdot \mathbf{v} \quad \text{in } \Omega,$$
$$\nabla \times \mathbf{w} = \tilde{g} + f_2 - \nabla \times \mathbf{v} \quad \text{in } \Omega,$$
$$\mathbf{n} \cdot \mathbf{w} = 0 \quad \text{on } \Gamma.$$

Letting $\tilde{\mathbf{u}} = \mathbf{w} + \mathbf{v}$, we have found a $(\tilde{\mathbf{u}}, \tilde{f}, \tilde{g})$ satisfying (35). Now, we need to verify that a $\tilde{\boldsymbol{\phi}}$ satisfying (35) exists. First, we construct a solution to the following linear ordinary differential equation

$$\frac{\partial \tilde{\boldsymbol{\phi}}}{\partial t} - \tilde{\boldsymbol{\phi}} \cdot \nabla_{\boldsymbol{\phi}} \mathbf{u}(\boldsymbol{\phi}) - \tilde{\mathbf{u}}(\boldsymbol{\phi}) = \mathbf{f}_3$$

along with initial condition $\tilde{\boldsymbol{\phi}}(0, \mathbf{x}) = \mathbf{f}_6(\mathbf{x})$ as

$$\begin{aligned}\tilde{\boldsymbol{\phi}}(t, \mathbf{x}) &= \mathbf{f}_6(\mathbf{x}) \\ &\quad + e^{K(t,\mathbf{x})} \int_0^t e^{-K(r,\mathbf{x})} (\tilde{\mathbf{u}}(\boldsymbol{\phi}(r, \mathbf{x})) + \mathbf{f}_3(r, \mathbf{x})) dr \\ &\quad + e^{K(t,\mathbf{x})} (\tilde{\mathbf{u}}(\boldsymbol{\phi}(0, \mathbf{x})) + \mathbf{f}_3(0, \mathbf{x})) t,\end{aligned}$$

where

$$K(k, \mathbf{x}) = \int_0^k \nabla_{\boldsymbol{\phi}} \mathbf{u}(\boldsymbol{\phi}(s, \mathbf{x}))^t ds + \nabla_{\boldsymbol{\phi}} \mathbf{u}(\boldsymbol{\phi}(0, \mathbf{x}))^t k.$$

Since $(\mathbf{u}, \boldsymbol{\phi}, f, g)$ is the optimal solution of (15), $\|f\|_1 + \|g\|_1$ is bounded and thus $\|\mathbf{u}\|_2$ is bounded. By Lemma 3, $|\nabla_{\boldsymbol{\phi}} \mathbf{u}(\boldsymbol{\phi})|$ is bounded *a.e.* in $(0, 1) \times \Omega$ as well as $e^{K(k,\mathbf{x})}$ is bounded *a.e.* in Ω. Hence $\tilde{\boldsymbol{\phi}} \in [L^2((0, 1); \Omega)]^2$. $\partial \tilde{\boldsymbol{\phi}} / \partial t \in [L^2((0, 1); \Omega)]^2$ follows easily. □

In the next theorem, we prove that a suitable Lagrange multiplier exists.

Theorem 2 *Let $(\mathbf{u}, \boldsymbol{\phi}, f, g) \in V$ denote an optimal solution to (15). Then, there exists a nonzero Lagrange multiplier $\boldsymbol{\Theta} \in W^*$ such that*

$$\begin{aligned}&J'(\boldsymbol{\phi}|_{t=1}, f, g) \cdot (\hat{\mathbf{u}}, \hat{\boldsymbol{\phi}}, \hat{f}, \hat{g}) \\ &\quad + \left\langle M'(\mathbf{u}, \boldsymbol{\phi}, f, g) \cdot (\hat{\mathbf{u}}, \hat{\boldsymbol{\phi}}, \hat{f}, \hat{g}), \boldsymbol{\Theta} \right\rangle = 0\end{aligned}$$

for any $(\hat{\mathbf{u}}, \hat{\boldsymbol{\phi}}, \hat{f}, \hat{g}) \in V$.

Proof Consider the nonlinear operator $N : V \to \mathbb{R} \times W$ defined by

$$N(\hat{\mathbf{u}}, \hat{\boldsymbol{\phi}}, \hat{f}, \hat{g}) = \begin{pmatrix} J(\hat{\boldsymbol{\phi}}|_{t=1}, \hat{f}, \hat{g}) - J(\boldsymbol{\phi}|_{t=1}, f, g) \\ M(\hat{\mathbf{u}}, \hat{\boldsymbol{\phi}}, \hat{f}, \hat{g}) \end{pmatrix},$$

where J and M are defined as in (5) and (7), respectively. Then the Frechét derivative $N'(\mathbf{u}, \boldsymbol{\phi}, f, g)$ is defined as

$$N'(\mathbf{u}, \boldsymbol{\phi}, f, g) \cdot (\hat{\mathbf{u}}, \hat{\boldsymbol{\phi}}, \hat{f}, \hat{g}) = (a, \mathbf{F})$$

for $(\hat{\mathbf{u}}, \hat{\boldsymbol{\phi}}, \hat{f}, \hat{g}) \in V$ and $(a, \mathbf{F}) \in \mathbb{R} \times W$ if and only if

$$\begin{aligned}&\left\langle \nabla \mathrm{T}(\boldsymbol{\phi}|_{t=1})(\mathrm{T}(\boldsymbol{\phi}|_{t=1}) - \mathrm{R}), \hat{\boldsymbol{\phi}} \right\rangle + \alpha_{f_0} \left\langle f, \hat{f} \right\rangle \\ &\quad + \alpha_{f_1} \left\langle \nabla f, \nabla \hat{f} \right\rangle + \alpha_{g_0} \langle g, \hat{g} \rangle + \alpha_{g_1} \langle \nabla g, \nabla \hat{g} \rangle = a\end{aligned}$$

and

$$M'(\mathbf{u}, \boldsymbol{\phi}, f, g) \cdot (\hat{\mathbf{u}}, \hat{\boldsymbol{\phi}}, \hat{f}, \hat{g}) = \mathbf{F}.$$

The operator $M'(\mathbf{u}, \boldsymbol{\phi}, f, g)$ is onto W by Lemma 4 and therefore has a closed range in W. Also, we know that $M'(\mathbf{u}, \boldsymbol{\phi}, f, g)$ is a linear operator from V to W. Thus, the kernel of $M'(\mathbf{u}, \boldsymbol{\phi}, f, g)$, $\ker(M'(\mathbf{u}, \boldsymbol{\phi}, f, g))$, is a closed subspace of V. Since $J'(\boldsymbol{\phi}|_{t=1}, f, g)$ acting on the kernel of $M'(\mathbf{u}, \boldsymbol{\phi}, f, g)$ is either identically zero or onto $\mathbb{R}$ (this follows from the obvious result that whenever f is a linear functional on a Banach space X, then either $f \equiv 0$ or the range of f, $\mathrm{ran}(f)$, is $\mathbb{R}$), $J'(\boldsymbol{\phi}|_{t=1}, f, g)$ acting on the kernel of $M'(\mathbf{u}, \boldsymbol{\phi}, f, g)$ 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, the $C : X \to Y \times Z$ defined by $Cx = (Ax, Bx)$ has a closed range in $Y \times Z$. Therefore, $N'(\mathbf{u}, \boldsymbol{\phi}, f, g)$ has a closed range in $\mathbb{R} \times W$.

Now, we suppose that $N'(\mathbf{u}, \boldsymbol{\phi}, f, g)$ is onto. Then there exists a $(\tilde{\mathbf{u}}, \tilde{\boldsymbol{\phi}}, \tilde{f}, \tilde{g}) \in A_{ad}$ satisfying

$$J(\tilde{\boldsymbol{\phi}}|_{t=1}, \tilde{f}, \tilde{g}) < J(\boldsymbol{\phi}|_{t=1}, f, g),$$

where $J((\boldsymbol{\phi} - \tilde{\boldsymbol{\phi}})|_{t=1}, f - \tilde{f}, g - \tilde{g}) < \epsilon$ with a small $\epsilon > 0$ which contradicts the hypothesis that $(\mathbf{u}, \boldsymbol{\phi}, f, g)$ is an optimal solution to (15).

To conclude, we use the Hahn-Banach theorem. Since $ran(N'(\mathbf{u}, \boldsymbol{\phi}, f, g))$ is closed in $\mathbb{R} \times W$, for any $(a, \mathbf{F}) \in ran(N'(\mathbf{u}, \boldsymbol{\phi}, f, g))$, there exists a nonzero $\boldsymbol{\Theta} \in (\mathbb{R} \times W)^*$ satisfying

$$\langle (a, \mathbf{F}), (\tilde{a}, \boldsymbol{\Theta}) \rangle = 0.$$

Suppose $\tilde{a} = 0$; then $\langle \mathbf{F}, \boldsymbol{\Theta} \rangle = 0$ for all $\mathbf{F}$. Therefore, $\tilde{a} \neq 0$. Now, without loss of generality, we can set $\tilde{a} = -1$ and therefore the result the theorem holds. □

4 The Optimality System

We showed the existence of optimal solutions and of a suitable Lagrange multiplier to turn the constrained optimization problem into an unconstrained optimization problem. In this section, we derive the optimality system from the formal calculation of variations of a Lagrangian functional. Let the optimal state $(\mathbf{u}, \boldsymbol{\phi}) \in [H^2(\Omega)]^2 \times [H^1((0,1); L^2(\Omega))]^2$, the optimal control $(f, g) \in H^1(\Omega) \times H^1(\Omega)$, and the corresponding Lagrange multiplier $(\xi, \eta, \boldsymbol{\psi}, \sigma, \nu, \boldsymbol{\mu}) \in W^*$, where

$$W^* = H^{-1}(\Omega) \times H^{-1}(\Omega) \times [L^2((0,1) \times \Omega)]^2 \times \mathbb{R} \times H^{-\frac{3}{2}}(\Gamma) \times [L^2(\Omega)]^2.$$

Define the penalized Lagrangian functional

$$\begin{aligned} &L(\mathbf{u}, \boldsymbol{\phi}, f, g; \xi, \eta, \boldsymbol{\psi}, \sigma, \nu, \boldsymbol{\mu}) \\ &\quad = J(\boldsymbol{\phi}|_{t=1}, f, g) - \int_\Omega (\nabla \cdot \mathbf{u} - f + 1)\xi d\mathbf{x} \\ &\qquad - \int_\Omega (\nabla \times \mathbf{u} - g)\eta d\mathbf{x} \\ &\qquad - \int_\Omega \int_0^1 \left(\frac{\partial \boldsymbol{\phi}}{\partial t} - \mathbf{u}(\boldsymbol{\phi})\right) \cdot \boldsymbol{\psi} dt d\mathbf{x} - \sigma \int_\Omega (f-1) d\mathbf{x} \\ &\qquad - \int_\Gamma (\mathbf{n} \cdot \mathbf{u}) \nu ds - \int_\Omega (\boldsymbol{\phi}(0, \mathbf{x}) - \boldsymbol{\phi}_0(\mathbf{x})) \cdot \boldsymbol{\mu} d\mathbf{x}, \end{aligned}$$

where $\alpha_{f_0}, \alpha_{f_1}, \alpha_{g_0}$, and α_{g_1} in the functional J are penalty parameters and β is the barrier parameter.

Variations of the Lagrangian with respect to the Lagrange multipliers recover the state system (7) in weak formulation, that is,

$$\begin{cases} \langle \nabla \cdot \mathbf{u}, \xi \rangle = \langle f - 1, \xi \rangle & \forall \xi \in H^{-1}(\Omega), \\ \langle \nabla \times \mathbf{u}, \eta \rangle = \langle g, \eta \rangle & \forall \eta \in H^{-1}(\Omega), \\ \langle \mathbf{n} \cdot \mathbf{u}, \nu \rangle_\Gamma = 0 & \forall \nu \in H^{-\frac{3}{2}}(\Gamma), \end{cases} \tag{37}$$

$$\begin{cases} \langle \frac{\partial \boldsymbol{\phi}}{\partial t} - \mathbf{u}(\boldsymbol{\phi}), \boldsymbol{\psi} \rangle = 0 & \forall \boldsymbol{\psi} \in [L^2((0,1) \times \Omega)]^2, \\ \langle \boldsymbol{\phi}(0, \mathbf{x}), \boldsymbol{\mu} \rangle = \langle \boldsymbol{\phi}_0(\mathbf{x}), \boldsymbol{\mu} \rangle & \forall \boldsymbol{\mu} \in [L^2(\Omega)]^2, \end{cases} \tag{38}$$

and

$$\sigma \int_\Omega \big(f(\mathbf{x}) - 1\big) d\mathbf{x} = 0 \quad \text{for all } \sigma \in \mathbb{R}. \tag{39}$$

Variations with respect to the state variables $\boldsymbol{\phi}$ and $\mathbf{u}$ yield the weak formulation of the adjoint equations

$$\begin{cases} \langle \frac{\partial \boldsymbol{\psi}}{\partial t} + \nabla_{\boldsymbol{\phi}} \mathbf{u}(\boldsymbol{\phi}) \boldsymbol{\psi}, \boldsymbol{\phi} \rangle = 0 \quad \forall \boldsymbol{\phi} \in [H^1((0,1); L^2(\Omega))]^2, \\ \langle \boldsymbol{\psi}(0, \mathbf{x}) - \boldsymbol{\mu}, \boldsymbol{\phi}(0, \mathbf{x}) \rangle = 0 \quad \forall \boldsymbol{\phi}(0, \mathbf{x}) \in [L^2(\Omega)]^2, \\ \langle \boldsymbol{\psi}(1, \mathbf{x}) - (T(\boldsymbol{\phi}(1, \mathbf{x})) \\ \quad - R(\mathbf{x})) \nabla_{\boldsymbol{\phi}} T(\boldsymbol{\phi}(1, \mathbf{x})), \boldsymbol{\phi}(1, \mathbf{x}) \rangle = 0 \\ \quad \forall \boldsymbol{\phi}(1, \mathbf{x}) \in [L^2(\Omega)]^2 \end{cases} \tag{40}$$

and

$$\begin{cases} \langle \nabla^\perp \eta - \nabla \xi, \mathbf{u} \rangle = \langle \int_0^1 |\nabla \boldsymbol{\phi}^{-1}(t, \mathbf{x})| \boldsymbol{\psi}(t, \boldsymbol{\phi}^{-1}(t, \mathbf{x})) dt, \mathbf{u} \rangle, \\ \int_\Gamma (\tau \cdot \mathbf{u}) \eta ds = 0, \\ \int_\Gamma (\mathbf{n} \cdot \mathbf{u}) \nu ds = 0 \quad \forall \mathbf{u} \in [H^2(\Omega)]^2, \end{cases} \tag{41}$$

where τ is the unit tangential vector on Γ. In (40) and (41), the equations $\boldsymbol{\mu} = \boldsymbol{\psi}(0, \mathbf{x})$ and $\nu = 0$ serve no role in defining $\boldsymbol{\psi}$, ξ, and η and merely serve to define the Lagrange multipliers $\boldsymbol{\mu}$ and ν, respectively. Thus, they may be omitted to yield a reduced set of adjoint equations.

We can construct a solution $\boldsymbol{\psi} \in [L^2((0,1) \times \Omega)]^2$ of (40) by

$$\boldsymbol{\psi}(t, \mathbf{x}) = e^{\kappa(t, \mathbf{x})} \boldsymbol{\psi}(1, \mathbf{x}),$$

where

$$\kappa(t, \mathbf{x}) = \int_t^1 \nabla_{\boldsymbol{\phi}} \mathbf{u}(\boldsymbol{\phi}(s, \mathbf{x})) ds + \nabla_{\boldsymbol{\phi}} \mathbf{u}(\boldsymbol{\phi}(1, \mathbf{x}))(1 - t).$$

Once $\boldsymbol{\psi}$ is determined, we can solve (41) for ξ and η since the right-hand side is known. In the following lemma, we show $|\nabla \boldsymbol{\phi}^{-1}(t, \mathbf{x})| \boldsymbol{\psi}(t, \boldsymbol{\phi}^{-1}(t, \mathbf{x}))$ is in $[L^2((0,1) \times \Omega)]^2$.

Lemma 5 *Let* $(\mathbf{u}, \boldsymbol{\phi}, f, g) \in V$ *be an optimal solution to* (15) *and* $\boldsymbol{\psi} \in [L^2((0,1) \times \Omega)]^2$. *Then,*

$$|\nabla \boldsymbol{\phi}^{-1}(t, \mathbf{x})| \boldsymbol{\psi}(t, \boldsymbol{\phi}^{-1}(t, \mathbf{x})) \in [L^2((0,1) \times \Omega)]^2.$$

Proof By the change of variables $(t, \mathbf{y}) = (t, \boldsymbol{\phi}^{-1}(t, \mathbf{x}))$ and $dt d\mathbf{y} = |\nabla \boldsymbol{\phi}^{-1}(t, \mathbf{x})| dt d\mathbf{x}$, we have

$$\begin{aligned} &\int_\Omega \int_0^1 \Big(|\nabla \boldsymbol{\phi}^{-1}(t, \mathbf{x})| \boldsymbol{\psi}(t, \boldsymbol{\phi}^{-1}(t, \mathbf{x}))\Big)^2 dt d\mathbf{x} \\ &\quad = \int_\Omega \int_0^1 |\nabla \boldsymbol{\phi}^{-1}(t, \mathbf{x})| (\boldsymbol{\psi}(t, \mathbf{y}))^2 dt d\mathbf{y} \\ &\quad = \int_\Omega \int_0^1 \frac{1}{|\nabla \boldsymbol{\phi}(t, \mathbf{y})|} (\boldsymbol{\psi}(t, \mathbf{y}))^2 dt d\mathbf{y}. \end{aligned} \tag{42}$$

Here, $\boldsymbol{\phi}$ satisfies

$$\frac{\partial \boldsymbol{\phi}}{\partial t} = \mathbf{u}(\boldsymbol{\phi}(t, \mathbf{x})).$$

Recall that if $dH(t)/dt = A(t)H(t)$, then $d(\det H(t))/dt = \operatorname{Tr} A(t) \det H(t)$ with $\operatorname{Tr} A(t)$ the trace of the matrix $A(t)$. Thus, we have

$$\begin{aligned} \frac{\partial \det \nabla \boldsymbol{\phi}}{\partial t} &= \operatorname{Tr} \nabla_{\boldsymbol{\phi}} \mathbf{u}(\boldsymbol{\phi}) \det \nabla \boldsymbol{\phi} = \nabla \cdot_{\boldsymbol{\phi}} \mathbf{u}(\boldsymbol{\phi}) \det \nabla \boldsymbol{\phi} \\ &= (f(\boldsymbol{\phi}) - 1) \det \nabla \boldsymbol{\phi} \end{aligned}$$

which, with the property $f > 0$, yields

$$\det \nabla \boldsymbol{\phi} = e^{\int_0^1 f(\boldsymbol{\phi})ds + (f(\mathbf{x})-1)t} \geq e^{-1}.$$

Therefore,

$$\int_\Omega \int_I \frac{1}{|\nabla \boldsymbol{\phi}(t,\mathbf{y})|} (\boldsymbol{\psi}(t,\mathbf{y}))^2 dt d\mathbf{y} \leq e \int_\Omega \int_I \boldsymbol{\psi}(t,\mathbf{y})^2 dt d\mathbf{y}.$$

Since $\boldsymbol{\psi} \in [L^2((0,1)\times\Omega)]^2$, $|\nabla \boldsymbol{\phi}^{-1}(t,\mathbf{x})|\boldsymbol{\psi}(t,\boldsymbol{\phi}^{-1}(t,\mathbf{x})) \in [L^2((0,1)\times\Omega)]^2$. □

From the above lemma, it is easy to see that the Lagrange multipliers η and ξ in fact satisfy

$$\eta, \xi \in L^2((0,1); H^1(\Omega)).$$

Variations of the Lagrangian with respect to the control f and g yield

$$\begin{cases} \langle -\alpha_{f_1}\Delta f + \alpha_{f_0} f - \frac{\beta}{f}, \tilde{f}\rangle = \langle \sigma - \xi, \tilde{f}\rangle & \forall \tilde{f} \in H^1(\Omega), \\ \int_\Gamma (\mathbf{n}\cdot\nabla f)\tilde{f} = 0 & \forall \tilde{f} \in H^1(\Omega), \\ \langle -\alpha_{g_1}\Delta g + \alpha_{g_0} g, \tilde{g}\rangle = \langle -\eta, \tilde{g}\rangle & \forall \tilde{g} \in H^1(\Omega), \\ \int_\Gamma (\mathbf{n}\cdot\nabla g)\tilde{g} = 0 & \forall \tilde{g} \in H^1(\Omega). \end{cases} \tag{43}$$

Note that one can enforce (39) by setting

$$\sigma = \int_\Omega \left(\xi - \frac{\beta}{f}\right) dx. \tag{44}$$

The above equations (37), (38), (39), (40), (41), and (43) are weak formulations of the following set of partial differential equations:

$$\begin{cases} \nabla\cdot\mathbf{u}(\mathbf{x}) = f(\mathbf{x}) - 1 & \text{in } \Omega, \\ \nabla\times\mathbf{u}(\mathbf{x}) = g(\mathbf{x}) & \text{in } \Omega, \\ \mathbf{n}\cdot\mathbf{u}(\mathbf{x}) = 0 & \text{on } \Gamma, \end{cases} \tag{45}$$

$$\begin{cases} \frac{\partial \boldsymbol{\phi}(t,\mathbf{x})}{\partial t} = \mathbf{u}(\boldsymbol{\phi}(t,\mathbf{x})) & \text{in } (0,1]\times\Omega, \\ \boldsymbol{\phi}(0,\mathbf{x}) = \boldsymbol{\phi}_0(\mathbf{x}) & \text{in } \Omega, \end{cases} \tag{46}$$

$$\int_\Omega \big(f(\mathbf{x}) - 1\big) d\mathbf{x} = 0, \tag{47}$$

and

$$\begin{cases} \frac{\partial \boldsymbol{\psi}}{\partial t} + \nabla_{\boldsymbol{\phi}}\mathbf{u}(\boldsymbol{\phi})\boldsymbol{\psi} = 0 & \text{in } (0,1)\times\Omega, \\ \boldsymbol{\psi}(1,\mathbf{x}) = (T(\boldsymbol{\phi}(1,\mathbf{x})) - R(\mathbf{x})) & \\ \qquad\qquad \times \nabla_{\boldsymbol{\phi}} T(\boldsymbol{\phi}(1,\mathbf{x})) & \text{in } \Omega, \\ \boldsymbol{\mu} = \boldsymbol{\psi}(0,\mathbf{x}) & \text{in } \Omega, \end{cases} \tag{48}$$

$$\begin{cases} \nabla^\perp\eta - \nabla\xi = \int_0^1 |\nabla\boldsymbol{\phi}^{-1}(t,\mathbf{x})| & \\ \qquad\qquad \times \boldsymbol{\psi}(t,\boldsymbol{\phi}^{-1}(t,\mathbf{x}))dt & \text{in } \Omega, \\ \eta = 0 & \text{on } \Gamma, \\ \nu = 0 & \text{on } \Gamma, \end{cases} \tag{49}$$

and

$$\begin{cases} -\alpha_{f_1}\Delta f + \alpha_{f_0} f - \frac{\beta}{f} = \sigma - \xi & \text{in } \Omega, \\ \mathbf{n}\cdot\nabla f = 0 & \text{on } \Gamma, \\ -\alpha_{g_1}\Delta g + \alpha_{g_0} g = -\eta & \text{in } \Omega, \\ \mathbf{n}\cdot\nabla g = 0 & \text{on } \Gamma, \end{cases} \tag{50}$$

respectively.

The optimality system is thus given by the state equations (45) and (46), the adjoint or co-state equations (48) and (49) with the third equation in each omitted, the optimality conditions (50), and (44). The optimality system is fully coupled. However, as briefly mentioned in the introduction, it can be shown that if one guesses control functions $f(\mathbf{x})$ and $g(\mathbf{x})$, then one can solve sequentially the uncoupled problems (45) for $\mathbf{u}(\mathbf{x})$, (46) for $\boldsymbol{\phi}(t,\mathbf{x})$, (48) for $\boldsymbol{\psi}(t,\mathbf{x})$, (49) for $\xi(\mathbf{x})$ and $\eta(\mathbf{x})$, and (44) for σ and then use (50) to update f and g. It can also be shown that this simple iterative process defines a steepest descent method. Of course, more sophisticated uncoupling strategies can be defined. Further details can be found in the upcoming paper where the numerical analysis of finite element approximations of the optimality system is studied and the results of several computational experiments are provided.

In the following section, we show some preliminary numerical results. Again, we omit the details about the numerical aspects since they are too long to describe here and will appear in the upcoming paper.

5 Numerical Results

Here, we present several examples of the application of our approach to image registration. In order to solve the optimality system (45)–(50), we use and iterative gradient algorithm; see [11] for the details about the gradient algorithm. A least-squares finite element method is used to solve the state equation (45), adjoint equation (49), and the optimality condition (50). The backward-Euler scheme is used for the time discretization of the ordinary differential equations (46) and (48) to obtain the updated $\boldsymbol{\phi}$ and $\boldsymbol{\psi}$, respectively.

5.1 Example 1

Let $\Omega = [0,1]\times[0,1]$. The reference and template images are given as

$$\mathrm{R}(\mathbf{x}) = \begin{cases} 30 & \text{if } \mathbf{x} \in [0.25, 0.5]\times[0.5, 0.75], \\ 10 & \text{otherwise} \end{cases}$$

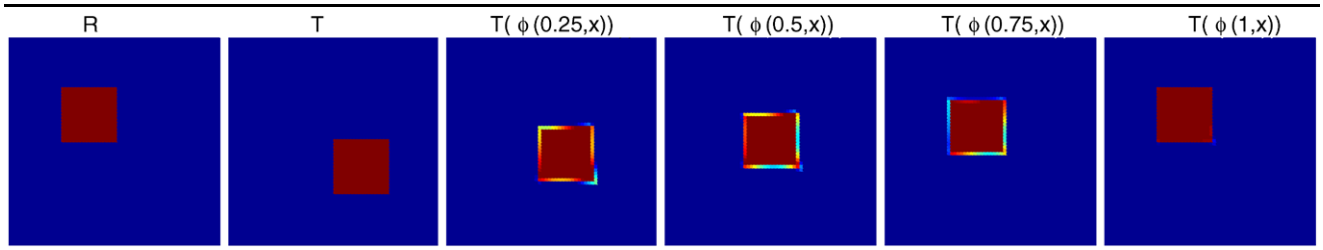


Fig. 1 *From left to right*: $\mathrm{R}(\mathbf{x})$ and $\mathrm{T}(\boldsymbol{\phi}(t,\mathbf{x}))$ at $t = 0, 0.25, 0.5, 0.75, 1$ for Example 1

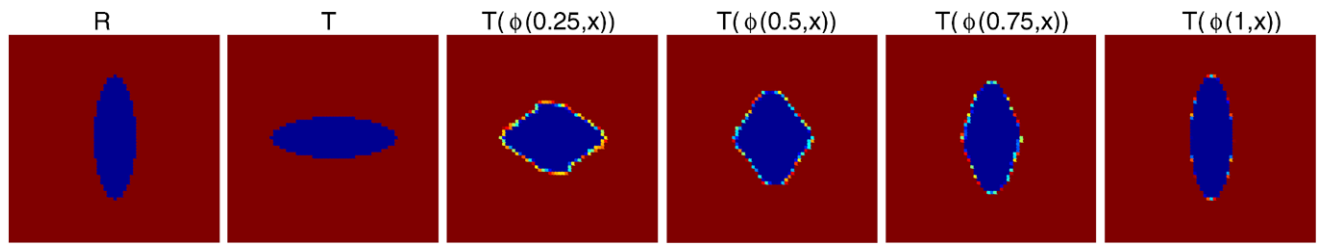


Fig. 2 *From left to right*: $\mathrm{R}(\mathbf{x})$ and $\mathrm{T}(\boldsymbol{\phi}(t,\mathbf{x}))$ at $t = 0, 0.25, 0.5, 0.75, 1$ for Example 2

Table 1 L^2-differences between the converged template image and the reference image for the translation of a square

$h = \Delta t =$	1/16	1/32	1/64
$\lVert \mathrm{T}(\boldsymbol{\phi}^h(1,\mathbf{x})) - \mathrm{R} \rVert$	0.195	0.072	0.037

Table 2 L^2-differences between the converged template image and the reference image for the rotation of an ellipse

$h = \Delta t =$	1/16	1/32	1/64
$\lVert \mathrm{T}(\boldsymbol{\phi}^h(1,\mathbf{x})) - \mathrm{R} \rVert$	0.680	0.447	0.239

and

$$\mathrm{T}(\mathbf{x}) = \begin{cases} 30 & \text{if } \mathbf{x} \in [0.5, 0.75] \times [0.25, 0.5], \\ 10 & \text{otherwise}, \end{cases}$$

respectively.

The penalty parameters are set to $\alpha_{f_0} = \alpha_{f_1} = \alpha_{g_0} = \alpha_{g_1} = 0.005$. For computational simplicity, we choose the barrier parameter $\beta = 0$ and force the values of f to be zero wherever, on each iteration, they are negative. Figure 1 shows, for the mesh size $h = 1/64$ and time step $\Delta t = 1/64$, plots of $\mathrm{T}(\boldsymbol{\phi}(t,\mathbf{x}))$ at $t = 0$, 0.125, 0.25, 0.5, 0.75 and 1 and the plot of $\mathrm{R}(\mathbf{x})$. The L^2-difference of $\mathrm{T}(\boldsymbol{\phi}(1,\mathbf{x}))$ and R, $\lVert \mathrm{T}(\boldsymbol{\phi}^h(1,\mathbf{x})) - \mathrm{R} \rVert$ is reduced to 0.037 by the gradient method compared to the starting value $\lVert \mathrm{T} - \mathrm{R} \rVert = 7.1$. The gradient iteration converges very slowly, therefore implementing faster iteration schemes will be part our future efforts. Table 1 shows the L^2-differences between the transformed template and reference images as the mesh size and time step change. Detailed analyses of approximation errors will be given in a follow-up paper.

5.2 Example 2

We next test the rotation of an ellipse. Again, let $\Omega = [0,1] \times [0,1]$. The reference and template images are given as

$$\mathrm{R}(\mathbf{x}) = \begin{cases} 1 & \text{if } (x-0.5)^2/0.01 + (y-0.5)^2/0.09 < 1, \\ 25 & \text{otherwise} \end{cases}$$

and

$$\mathrm{T}(\mathbf{x}) = \begin{cases} 1 & \text{if } (x-0.5)^2/0.09 + (y-0.5)^2/0.01 < 1, \\ 25 & \text{otherwise}, \end{cases}$$

respectively, where, for the starting template image, $\lVert \mathrm{T} - \mathrm{R} \rVert = 9.25$. Figure 2 provides plots of results using a mesh size $h = 1/64$ and time step $\Delta t = 1/64$. Again, we the parameters were set to $\alpha_{f_0} = \alpha_{f_1} = \alpha_{g_0} = \alpha_{g_1} = 0.0005$ and $\beta = 0$. In Table 2, the values of $\lVert \mathrm{T}(\boldsymbol{\phi}^h(1,\mathbf{x})) - \mathrm{R}(\mathbf{x}) \rVert$ as the mesh size and time step change are shown.

5.3 Example 3

Our next example is a transformation from a circle to the letter C. The reference and template images are given as 65×65 data with values

$$\mathrm{R}(\mathbf{x}) = \begin{cases} 50 & \text{within the C}, \\ 1 & \text{otherwise} \end{cases}$$

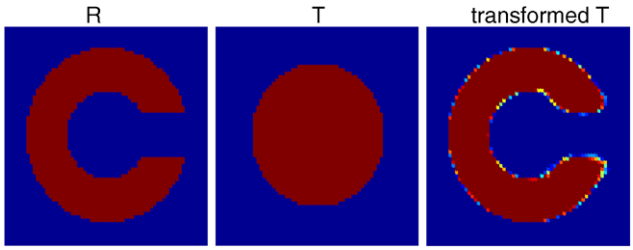


Fig. 3 *From left to right*: R and starting and transformed T for Example 3

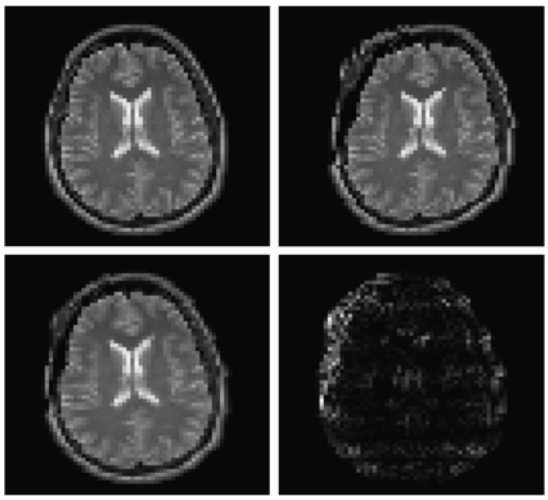


Fig. 4 *Left to right* and *top to bottom*: R, T, transformed T, and $|\mathrm{T}(\boldsymbol{\phi}(1,\mathbf{x})) - \mathrm{R}|$ for Example 4

and

$$\mathrm{T}(\mathbf{x}) = \begin{cases} 50 & \text{within the circle,} \\ 1 & \text{otherwise;} \end{cases}$$

we have that $\|\mathrm{T} - \mathrm{R}\| = 26.94$. In Fig. 3, the reference, template, and transformed template images with $\alpha_{f_0} = \alpha_{f_1} = \alpha_{g_0} = \alpha_{g_1} = 0.0001$, $\beta = 0$, and $\Delta t = 1/64$ are shown. After the iterations converge, the L^2-difference between the transformed template and reference images is reduced to 2.409.

5.4 Example 4

Now, we demonstrate the registration of an MRI image of a brain. The reference image, R, and the template image, T, are given as 65×65 data. The parameters $\alpha_{f_0} = \alpha_{f_1} = \alpha_{g_0} = \alpha_{g_1} = 0.0005$, $\beta = 0$ and $\Delta t = 1/64$ are fixed. In Fig. 4, the left and right images in the first row are the given reference and template images, respectively. The left image in second row is the transformed template image and the right image shows the absolute value of the difference $\mathrm{T}(\boldsymbol{\phi}(1,\mathbf{x})) - \mathrm{R}$. The starting L^2-difference, $\|\mathrm{T} - \mathrm{R}\|$, is 2.255 and, after optimization, the difference between the transformed template and reference images is reduced to $\|\mathrm{T}(\boldsymbol{\phi}(1,\mathbf{x})) - \mathrm{R}\| = 0.48$.

6 Conclusion

An important goal of the image registration problem is to find a mapping that transforms a given template image to one similar to a reference image. In this paper, to find such a mapping, we describe an optimal control problem based on the grid deformation method and then turn the constrained minimization problem into an unconstrained problem by using the Lagrange multiplier rule. To our knowledge, this is the first attempt to combine the grid deformation method and Lagrange multiplier rule to solve the image registration problem. We prove that solutions of the optimal control problem exist and that the Lagrange multiplier rule can be applied to redefine the problem into an unconstrained optimization problem. We then derive an optimality system from which optimal solutions can be determined. Several preliminary computational examples are provided.

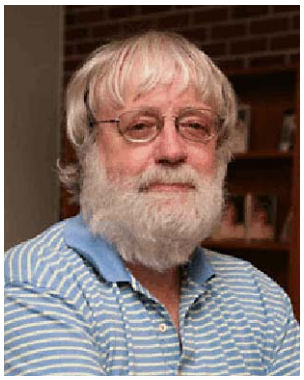
Although we only treated the two-dimensional setting, we anticipate that the methods and theories in this paper are extendible to three-dimensions. A future paper will deal with the development and analysis of finite element discretizations of the optimality system as well as with more extensive computational experiments.

References

1. Adams, R.A.: Sobolev Spaces. Pure and Applied Mathematics, vol. 65. Academic Press, New York (1975)
2. Amit, Y.: A nonlinear variational problem for image matching. SIAM J. Sci. Comput. **15**(1), 207–224 (1994)
3. Antoine Maintz, J.B., Vierge, M.A.: A survey of medical image registration. Med. Image Anal. **2**(1), 1–36 (1998)
4. Bertsekas, D.P.: Constrained Optimization and Lagrange Multiplier Methods. Computer Science and Applied Mathematics. Academic Press, New York (1982)
5. Birkhoff, G., Rota, G.C.: Ordinary Differential Equations, 3rd edn. Wiley, New York (1978)
6. Bochev, P., Liao, G., dela Pena, G.: Analysis and computation of adaptive moving grids by deformation. Numer. Methods Partial Differ. Equ. **12**(4), 489–506 (1996)
7. Christensen, G.E., Rabbitt, R.D., Miller, M.I.: Deformable templates using large deformation kinematics. IEEE Trans. Image Process. **5**(10), 1435–1447 (1996)
8. Droske, M., Ring, W.: A Mumford-Shah level-set approach for geometric image registration. SIAM J. Appl. Math. **66**(6), 2127–2148 (2006) (electronic)
9. Dupuis, P., Grenander, U., Miller, M.I.: Variational problems on flows of diffeomorphisms for image matching. Q. Appl. Math. **56**(3), 587–600 (1998)
10. Fischer, B., Modersitzki, J.: Curvature based image registration. J. Math. Imaging Vis. **18**(1), 81–85 (2003). Special issue on imaging science (Boston, MA, 2002)
11. Gunzburger, M.D.: Perspectives in Flow Control and Optimization. Advances in Design and Control, vol. 5. SIAM, Philadelphia (2003)
12. Haber, E., Heldmann, S., Modersitzki, J.: An octree method for parametric image registration. SIAM J. Sci. Comput. **29**(5), 2008–2023 (electronic)

13. Haber, E., Modersitzki, J.: Numerical methods for volume preserving image registration. Inverse Probl. **20**(5), 1621–1638 (2004)
14. Haber, E., Modersitzki, J.: A multilevel method for image registration. SIAM J. Sci. Comput. **27**(5), 1594–1607 (2006) (electronic)
15. Henn, S.: A full curvature based algorithm for image registration. J. Math. Imaging Vis. **24**(2), 195–208 (2006)
16. Henn, S., Witsch, K.: Iterative multigrid regularization techniques for image matching. SIAM J. Sci. Comput. **23**(4), 1077–1093 (2001) (electronic)
17. Liao, G., Pan, T.w., Su, J.: Numerical grid generator based on Moser's deformation method. Numer. Methods Partial Differ. Equ. **10**(1), 21–31 (1994)
18. Liao, G., Cai, X., Fleitas, D., Luo, X., Wang, J., Xue, J.: Optimal control approach to data set alignment. Appl. Math. Lett. **21**(9), 898–905 (2007)
19. McLachlan, R.I., Marsland, S.: Discrete mechanics and optimal control for image registration. ANZIAM J. **48**(C), C1–C16 (2007)
20. Modersitzki, J.: Numerical Methods for Image Registration Numerical Mathematics and Scientific Computation. Oxford University Press, New York (2004)
21. Rama Mohana Rao, M.: Ordinary Differential Equations. Affiliated East-West Press Pvt. Ltd., New Delhi (1980). Theory and applications, with a foreword by V. Lakshmikantham
22. Semper, B., Liao, G.: A moving grid finite-element method using grid deformation. Numer. Methods Partial Differ. Equ. **11**(6), 603–615 (1995)
23. Unal, G., Slabaugh, G.: Estimation of vector fields in unconstrained and inequality constrained variational problems for segmentation and registration. J. Math. Imaging Vis. **31**(1), 57–72 (2008)
24. Zeidler, E.: Nonlinear Functional Analysis and Its Applications, vol. II/A. Springer, New York (1990). Linear monotone operators, translated from the German by the author and Leo F. Boron

Eunjung Lee received bachelor's and master's degrees from Kyungpook National University in Republic of Korea and she received her Ph.D. from University of Colorado at Boulder in US in 2005. She is currently an assistant professor of department of computational science and engineering at Yonsei university in Seoul, Republic of Korea.



Max Gunzburger received his doctorate in 1969 from New York University, where he also earned his bachelor's and master's degrees. He began his career as a research scientist and assistant professor at New York University and followed that with research positions at the Naval Ordnance Laboratory and the Institute for Computer Applications in Science and Engineering. He was on the faculty of University of Tennessee, Carnegie Mellon University, Virginia Polytechnic Institute and State University, and Iowa State university. He is currently a Francis Eppes professor of scientific computing and a chair of department of scientific computing at Florida State university. He is a member of SIAM and has served as editor-in-chief of the SIAM Journal on Numerical Analysis and as chair of the Board of Trustees of SIAM. He was awarded the W.T. and Idelia Reid Prize in Mathematics in 2008. In 2009, he has been named as one of the charter fellows of SIAM.