

Technical Report

TR-2004-001

Numerical methods for volume preserving image registration

by

Eldad Haber, Jan Modersitzki

MATHEMATICS AND COMPUTER SCIENCE

EMORY UNIVERSITY

Numerical methods for volume preserving image registration

Eldad Haber* Jan Modersitzki*

February 17, 2004

Abstract

Image registration techniques are used routinely in a variety of today's medical imaging diagnosis. Since the problem is ill-posed, one may like to add additional information about distortions. This applies, for example, to the registration of contrast-enhanced images, where variations of substructures are not related to patient motion but to contrast uptake. Here, one may only be interested in registrations which do not alter the volume of any substructure.

In this paper we discuss image registration techniques with a focus on volume preserving constraints. These constraints can reduce the non-uniqueness of the registration problem significantly. Our implementation is based on a constrained optimization formulation. Upon discretization, we obtain a large, discrete, highly nonlinear optimization problem and the necessary conditions for the solution form a discretized nonlinear partial differential equation. To solve the problem we use a variant of the Sequential Quadratic Programming method. Moreover, we present results on synthetic as well as on real-life data.

1 Introduction

Image registration is one of the fundamental tasks in today's image processing and in particular in medical imaging; see, e.g., [13] and references therein.

*Dept of Mathematics and Computer Science, Emory University, Atlanta GA 30322
{haber,modersit}@mathcs.emory.edu.

The objective of image registration is to make images which are taken at different times, from different perspectives, and/or from different devices to be more alike. Loosely, the goal of image registration is to find a “*reasonable*” deformation such that the “*distance*” between a reference image R and a deformed version of a template image T becomes small.

An application of particular clinical interest is the registration of pairs of images acquired before and after contrast administration; see, e.g., [30] and references therein. A typical example is depicted in Fig. 1. In this application, magnetic resonance images of a female breast are taken at different times (images from Bruce Daniel, Lucas Center for Magnetic Resonance Spectroscopy and Imaging, Stanford University). The first image shows an MRI section taken during the so-called wash-in phase of a radiopaque marker and the second image shows the analogous section during the so-called wash-out phase. A comparison of these two images indicates a suspicious region in the upper part of the images. This region can be detected easily if the images have been registered: tissue located at a certain position in the wash-in image is related to the tissue at the same position in the wash-out phase. Generally, however, a quantitative analysis is a delicate matter since observable differences are not only related to contrast uptake but also due to motion of the patient, like, for example, breathing or heart beat.

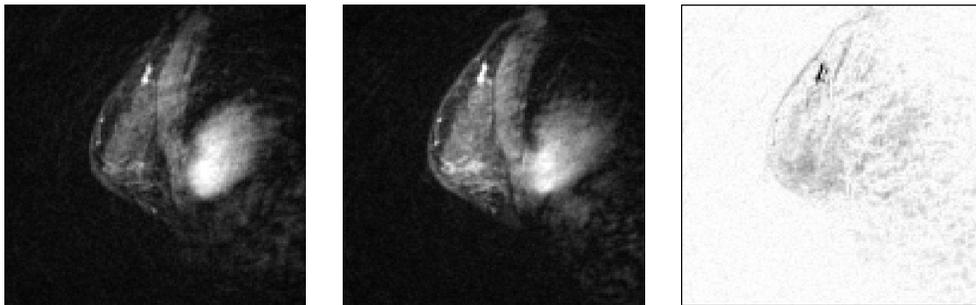


Figure 1: MRI’s of a female breast, LEFT: during the wash-in phase, MIDDLE: during the wash-out phase, and RIGHT: difference image.

As pointed out by Rohlfing et al. [30], there is a substantial difficulty with the registration of pre and post-contrast images. Bright regions seem to enlarge during the so-called wash-in phase. This enhancement is due to contrast uptake but not to movement of the patient. Fig. 2 illustrates an ideal situation. Without external information, it is impossible to answer whether

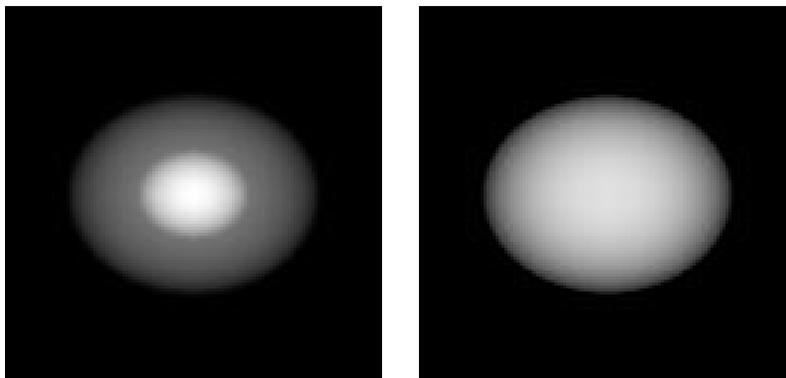


Figure 2: Synthetic images simulating contrast uptake, LEFT: wash-in, RIGHT: wash-out.

the white area has been enlarged or the grey area turned to white.

The idea is to restrict the set of feasible transformations a priori in a reasonable way. For this situation, we constrain the transformations to be volume preserving (VP). In contrast to [30], we use a variational setting. Therefore, we are not restricted to parametric and in particular B-spline based deformations. Also, we do not introduce an additional penalty term but use a constrained approach.

The VP approach is also connected to the so-called mass preserving (MP) registration which is related to the Monge-Kantorovich transport problem; see [19, 36]. From a mathematical point of view, the MP approach contains the VP approach (setting the density to 1). However, in the MP approach the images enter directly into the constraints, whereas in the VP they do not. In contrast to the MP approach, the VP approach aims to correct geometrical distortions but do not alter gray values. Another major difference between the above work and ours is the numerical treatment. For the MP approach, a Helmholtz decomposition is exploited, and, after eliminating the constraints, the necessary conditions for the optimization problem are solved using a steepest descent method. Here, we discretize the transformation and the constraints directly and use a Sequential Quadratic Programm to solve the discrete optimality conditions.

In this paper, we present a flexible constrained image registration approach. It has three main ingredients: a distance measure, a regularizer, and the constraints.

Our mathematical framework is general enough to handle a variety of distance measures, including the most popular ones, like those based on the sum of squared differences (SSD), mutual information (MI), or correlation, as long as a Gâteaux derivative exists; see, e.g., [29, 22, 20]. For presentation purposes, we explicitly discuss the approach only for the SSD measure.

Even with this additional constraints, however, image registration is an ill-posed problem, cf., e.g., [21, 26]. If one considers for example the registration of an image of a disc to a copy of this image, any rotation of the image gives a solution with respect to any reasonable distance measure. Note that a pure rotation is a rigid transformation and rigid transformations are volume preserving. For this reason, also the constrained image registration problem has to be regularized. The framework presented here is based on a general regularizer, too. Any regularizing term with a Gâteaux derivative can be used. This includes well-known choices, like, for example, the elastic [4, 6, 15, 9], the diffusion [23, 10], and the biharmonic [1, 12] regularizer. For ease of presentation, we emphasize on the most popular of these, the elastic registration.

Also our approach to the constraints is very general. However, since the implementation of the volume preserving constraints is not straightforward, we restrict ourselves to these constraints. It is very important to note that volume preservation constraints are pointwise differential constraints which after discretization apply to any pixel/voxel.

Finally, we suggest the use of a staggered grid discretization. This is a well-known, established, and often used technique in many fields, like for example fluid dynamics or electromagnetics. However, we are not aware of any image registration algorithm based on this discretization.

The paper is organized as follows. In Section 2 we present the continuous mathematical setup of the constrained registration problem and derive the continuous Euler-Lagrange equations. Although we do not use these equations in our numerical implementation, they give insight into the problem's structure and serve as a reference for the discrete analogs. In Section 3 we discuss the discretization of the problem. For readers from image processing, which might not be so familiar with staggered grids, we give a brief and formal introduction. Staggered grid discretization is well known to be stable when working with tightly coupled system of partial differential equations.

Our constraints are discretized using a finite volume discretization of each displaced pixel/voxel and we show that the resulting formula mimics

the continuous one. In Section 4 we discuss a numerical scheme for the optimization of the discretized image registration problem. In Section 5 we present numerical results and finally, in Section 6 we summarize the paper and discuss future work.

2 Mathematical setup

With $d \in \mathbb{N}$ we denote the spatial dimension of the given images $R, T : \mathbb{R}^d \rightarrow \mathbb{R}$ which are assumed to be sufficiently smooth. Thus, $T(\mathbf{x})$ gives a gray value at a spatial position \mathbf{x} . We assume that the supports of the images are contained in a bounded domain $\Omega :=]0, L[^d$, i.e. $R(\mathbf{x}) = T(\mathbf{x}) = 0$ for $\mathbf{x} \notin \Omega$.

Our goal is to find a “reasonable” deformation \mathbf{u} such that the “distance” between the reference image R and the deformed template image $T(\mathbf{x} + \mathbf{u}(\mathbf{x}))$ becomes small. Note that $\mathbf{u} = (u^1, \dots, u^d)^T : \mathbb{R}^d \rightarrow \mathbb{R}^d$. It is well-known that this problem is ill-posed and therefore needs to be regularized. In general it is common to use a Tikhonov-style regularization. A mathematical formulation of the regularized and constrained problem reads

$$\mathcal{D}[R, T; \mathbf{u}] + \alpha \mathcal{S}[\mathbf{u}] = \min \tag{1a}$$

$$\text{subject to } \mathcal{C}[\mathbf{u}](\mathbf{x}) = 0 \quad \text{for all } \mathbf{x} \in \Omega, \tag{1b}$$

where \mathcal{D} is some distance measure, \mathcal{S} is some regularization term, and \mathcal{C} are some constraints. Here, $\alpha > 0$ is a regularization parameter and compromises between similarity and regularity.

The three building blocks are discussed below. The constraints can be used to supply additional information about the registration problem. For example, in some applications it is of importance that particular points, like, e.g., anatomical landmarks, are in a precise one-to-one correspondence. With the setting $\mathcal{C}[u] = \ell^T + \mathbf{u}(\ell^T) - \ell^R$, one can guaranty the correspondence of the landmarks ℓ^R and ℓ^T ; see, e.g., [5, 11]. In this paper we focus on applications demanding for volume preserving (VP) transformations. However, one may set $\mathcal{C}[\mathbf{u}] \equiv 0$ to recover the unconstrained problem.

For the following analysis and numerics any choices and combinations of the building blocks \mathcal{D} , \mathcal{S} , and \mathcal{C} are feasible as long as they do have a Gâteaux derivative. For the purpose of this presentation, we restrict the discussion to the following choices.

Distance measures The distance measure used in this note is the so-called sum of squared differences (or L_2 norm)

$$\mathcal{D}[R, T; \mathbf{u}] = \frac{1}{2} \int_{\mathbb{R}^d} \left(T(\mathbf{x} + \mathbf{u}(\mathbf{x})) - R(\mathbf{x}) \right)^2 d\mathbf{x} \quad (2)$$

with Gâteaux derivative (cf., e.g., [26])

$$d_{\mathbf{u}, \mathbf{v}} \mathcal{D}[R, T; \mathbf{u}] = \int_{\mathbb{R}^d} \langle \mathbf{f}(\mathbf{x}, \mathbf{u}(\mathbf{x})), \mathbf{v}(\mathbf{x}) \rangle_{\mathbb{R}^d} d\mathbf{x}, \quad (3)$$

$$\mathbf{f}(\mathbf{x}, \mathbf{u}(\mathbf{x})) := (T(\mathbf{x} + \mathbf{u}(\mathbf{x})) - R(\mathbf{x})) \cdot \nabla T(\mathbf{x} + \mathbf{u}(\mathbf{x})), \quad \mathbf{x} \in \mathbb{R}^d. \quad (4)$$

Other distance measures like, e.g., those based on mutual information [7, 34], might be used as well.

Remark 1 *Note that due to the chain rule any differentiable distance functional based on R and $T(\mathbf{x} + \mathbf{u}(\mathbf{x}))$ has a factor $\nabla T(\mathbf{x} + \mathbf{u}(\mathbf{x}))$. Thus, since we assume R and T to be zero for $\mathbf{x} \notin \Omega$ also \mathbf{f} is zero for $\mathbf{x} \notin \Omega$. Therefore, the integration in Eqs. (2) and (3) reduces to an integration over the domain Ω , only.*

Regularizer In this work we consider the well-known elastic regularizer [4, 6, 9]. However, our formulation is flexible enough that we could use many other regularizers, like, e.g., the fluid [6, 3], diffusion [10], or curvature regularizers [12] or any combinations of these. Each of the above regularizers is based on a differential operator \mathcal{B} . Particularly,

$$\mathcal{S}[\mathbf{u}] = \int_{\Omega} \langle \mathcal{B}[\mathbf{u}], \mathcal{B}[\mathbf{u}] \rangle_{\mathbb{R}^d} d\mathbf{x}, \quad (5)$$

and therefore its Gâteaux derivative is

$$d_{\mathbf{u}, \mathbf{v}} \mathcal{S}[\mathbf{u}] = \int_{\Omega} \langle \mathcal{A}[\mathbf{u}](\mathbf{x}), \mathbf{v}(\mathbf{x}) \rangle_{\mathbb{R}^d} d\mathbf{x}, \quad \mathcal{A} = \mathcal{B}^* \mathcal{B}, \quad (6)$$

where, for ease of presentation, we assume natural boundary conditions on \mathbf{u} . For the elastic regularizer with Lamé constants λ and μ , we have

$$\mathcal{B}[\mathbf{u}] = \begin{pmatrix} \sqrt{\mu} & 0 \\ 0 & \sqrt{2\mu + \lambda} \end{pmatrix} \begin{pmatrix} \nabla \times \mathbf{u} \\ \nabla \cdot \mathbf{u} \end{pmatrix}, \quad \mathcal{A}[\mathbf{u}] = -\mu \Delta \mathbf{u} - (\mu + \lambda) \nabla \nabla \cdot \mathbf{u} \quad (7)$$

with $\nabla \cdot$ the divergence, $\nabla \times$ the curl, and Δ the Laplacian.

Constraints We require the transformation $\boldsymbol{\varphi}(\mathbf{x}) := \mathbf{x} + \mathbf{u}(\mathbf{x})$ to be volume preserving (see also [30]). The transformation $\boldsymbol{\varphi}$ is volume preserving if and only if for any domain V ,

$$\int_{\boldsymbol{\varphi}(V)} d\mathbf{x} = \int_V d\mathbf{x}, \quad \text{where } \boldsymbol{\varphi}(V) = \{\boldsymbol{\varphi}(\mathbf{x}) : \mathbf{x} \in V\}. \quad (8)$$

This constraint implies that not only the overall volume but also and most importantly the volume of each arbitrarily small subdomain is conserved. For a smooth transformation $\boldsymbol{\varphi}$, we use the transformation rule to derive the point wise constraints

$$1 = \det(\nabla \boldsymbol{\varphi}) = \det(I_d + \nabla \mathbf{u}) \quad (9)$$

$$\text{or } \mathcal{C}[\mathbf{u}] := \det(I_d + \nabla \mathbf{u}) - 1 = 0. \quad (10)$$

Here, $I_d \in \mathbb{R}^{d,d}$ denotes the d -by- d identity matrix and $\delta_{j,k}$ its (j, k) th entry.

The Gâteaux derivative of the volume preserving constraints is given by the following lemma.

Lemma 1 *Let \mathbf{v} be a suitable perturbation of \mathbf{u} . The Gâteaux derivative of \mathcal{C} (cf. Eq. (10)) is given by*

$$d_{\mathbf{u}, \mathbf{v}} \mathcal{C}[\mathbf{u}](\mathbf{x}) = \det(I_d + \nabla \mathbf{u}(\mathbf{x})) \langle (I_d + \nabla \mathbf{u}(\mathbf{x}))^{-\top}, \nabla \mathbf{v}(\mathbf{x}) \rangle_{\mathbb{R}^{d,d}}. \quad (11)$$

Proof: A computation gives

$$\begin{aligned} & d_{\mathbf{u}, \mathbf{v}} \det(I_d + \nabla \mathbf{u}) \\ &= \lim_{\varepsilon \rightarrow 0} (\det(I_d + \nabla \mathbf{u} + \varepsilon \nabla \mathbf{v}) - \det(I_d + \nabla \mathbf{u})) \\ &= \det(I_d + \nabla \mathbf{u}) \lim_{\varepsilon \rightarrow 0} \left(\det(I_d + \varepsilon \nabla \mathbf{v} (I_d + \nabla \mathbf{u})^{-1}) - 1 \right) \\ &= \det(I_d + \nabla \mathbf{u}) \cdot \text{trace} [\nabla \mathbf{v} (I_d + \nabla \mathbf{u})^{-1}] \\ &= \langle \det(I_d + \nabla \mathbf{u}) (I_d + \nabla \mathbf{u})^{-\top}, \nabla \mathbf{v} \rangle_{\mathbb{R}^{d,d}}. \quad \blacksquare \end{aligned}$$

Remark 2 *Applying Cramer's rule, it can be verified that the Gâteaux derivative of \mathcal{C} is a polynomial in $\partial_j u^k$, $j, k = 1, \dots, d$.*

Example 1 For $d = 2$ we have

$$\begin{aligned} \mathcal{C}[\mathbf{u}] &= \nabla \cdot \mathbf{u} + \partial_1 u_1 \partial_2 u_2 - \partial_2 u_1 \partial_1 u_2, \\ \det(I_2 + \nabla \mathbf{u})(I_2 + \nabla \mathbf{u})^{-\top} &= \begin{pmatrix} 1 + \partial_2 u_2 & -\partial_1 u_2 \\ -\partial_2 u_1 & 1 + \partial_1 u_1 \end{pmatrix}. \end{aligned}$$

We now investigate necessary conditions for a solution of the image registration problem (1). For computational purposes, we have to discretize either the optimization problem (1) or the resulting necessary conditions. In the following section we choose to discretize the optimization problem directly and therefore, the continuous conditions are not used directly in our numerical scheme. However, the discrete conditions have to mimic the continuous analogs and therefore we find it useful to study the latter.

Introducing the Lagrange multiplier $p : \mathbb{R}^d \rightarrow \mathbb{R}$, the Lagrangian of (1) is given by

$$\mathcal{L}[\mathbf{u}, p] = \mathcal{D}[R, T; \mathbf{u}] + \alpha \mathcal{S}[\mathbf{u}] + \int_{\Omega} \mathcal{C}[u](\mathbf{x}) \cdot p(\mathbf{x}) \, d\mathbf{x} \quad (12)$$

and the continuous Euler-Lagrange equations for (1) read

$$\begin{aligned} 0 &= d_{\mathbf{u}; \mathbf{v}} \mathcal{L}[\mathbf{u}, p] \\ &= d_{\mathbf{u}; \mathbf{v}} \mathcal{D}[R, T; \mathbf{u}] + \alpha d_{\mathbf{u}; \mathbf{v}} \mathcal{S}[\mathbf{u}] + \int_{\Omega} d_{\mathbf{u}; \mathbf{v}} \mathcal{C}[\mathbf{u}](\mathbf{x}) \cdot p(\mathbf{x}) \, d\mathbf{x}, \end{aligned} \quad (13a)$$

$$0 = d_{p; q} \mathcal{L}[\mathbf{u}, p] = \int_{\Omega} \mathcal{C}[\mathbf{u}](\mathbf{x}) \cdot q(\mathbf{x}) \, d\mathbf{x}, \quad (13b)$$

for any appropriate perturbations \mathbf{v} and q . Thus, for any $\mathbf{x} \in \Omega$, we have

$$\begin{aligned} 0 &= \mathbf{f}(\mathbf{x}, \mathbf{u}(\mathbf{x})) + \alpha \mathcal{A}[\mathbf{u}](\mathbf{x}) \\ &\quad - \nabla \cdot [\det(I_d + \nabla \mathbf{u}(\mathbf{x}))(I_d + \nabla \mathbf{u}(\mathbf{x}))^{-\top} \cdot p(\mathbf{x})] \end{aligned} \quad (14a)$$

$$0 = \det(I_d + \nabla \mathbf{u}(\mathbf{x})) - 1, \quad (14b)$$

where we imposed zero Dirichlet boundary conditions for the Lagrange multiplier p .

The system (14) presents a highly coupled system of nonlinear partial differential equations (PDE). The quantity \mathbf{f} in (14a), which depends nonlinearly on \mathbf{u} and the images, can be viewed as forces pushing the template towards similarity. The differential operator \mathcal{A} in (14b) is a linear, elliptic

operator. The last term in (14a) is related to the derivative of the constraints which also show up in (14b). For a simpler case, when the \mathbf{f} and \mathbf{A} in (14a) are replaced with the identity operator, existence and uniqueness of a solution can be shown [8, p. 324]. However, for the registration problem, it is not easy to show either existence or uniqueness of a solution of the PDE (14). For the purpose of this paper, we therefore assume existence of a solution and remark that proving its existence is a subject of further research.

3 Discretization

There are two approaches for the discretization of the PDE constrained optimization problem (1). In the first so-called *optimize-discretize* approach one forms the Lagrangian (12), then differentiates to obtain the continuous Euler-Lagrange equations (13), which are finally discretized and solved.

The second approach, that we use here, is the so-called *discretize-optimize* approach. Here, one directly discretizes the problem (1) and then solves a constrained optimization problem in a finite (but typically high) dimensional space. Note that we still use the fact that the discrete Euler-Lagrange equations are a discretization of some differential operators. The advantage of this approach is that we are able to use standard optimization methods for the solution of the problem.

Choosing a stable discretization method for an optimization problem with a differential constraint is a delicate matter. It is well known that such a discretization should fulfill the LBB conditions [2]. It is also well known that some seemingly good discretization methods do not fulfill this condition (see, e.g., [17] for an elaborate discussion for the Stokes system). Further complications arise in our case where we have differential operators such as the divergence and the curl. We would like to choose a conservative compact discretization scheme and this could be achieved by either mixed finite elements or by staggered grids. Staggered grids are very common for the stable discretization of fluid flow problems (see, e.g., [14]) where first order differential constraints are discretized and in electromagnetics (see, e.g., [35, 18]) where operators such as curl and divergence are discretized. In the context of fluid flows and electromagnetics it is well known that compact discretizations are crucial in order to obtain a stable linear system of equation and to avoid spurious modes; [24, 25, 17]. It is therefore most natural to choose such a discretization for our problem as well. Further investigation is

needed to show that the LBB conditions are fulfilled and this will be done in a subsequent paper.

Though staggered grids seem to be natural for the discretization of the registration problem on a regular grid, we are not aware of any registration scheme where this discretization is used. We therefore give a brief but formal description; see Section 3.1. For a more elaborate discussion, see, e.g., [18].

It is important to note that using staggered grids only short differences are used for the approximation of the derivatives $\partial_j u^k$ and therefore we do not introduce spurious modes by discretization; see, e.g., [35].

3.1 Staggered grid discretization

We assume that our discrete images have $m_1 \times \dots \times m_d$ pixels and, for ease of presentation, that each pixel is square with lengths h . We allow for half step indices. As usual in image processing, we identify pixels/voxels with cell centered grid points with are therefore labeled with full integers indices. The knots of the nodal, cell centered, the d face staggered, and (for dimension $d = 3$) the d edge staggered grids are collected in d -dimensional arrays as follows (see also Fig. 3 for an illustration),

$$\begin{aligned} X^n &= h(i_1 - \frac{1}{2}, \dots, i_d - \frac{1}{2})_{i_j = \frac{1}{2}, \dots, m_j + \frac{1}{2}}, \\ X^c &= h(i_1 - \frac{1}{2}, \dots, i_d - \frac{1}{2})_{i_j = 1, \dots, m_j}, \\ X^{f,j} &= h(i_1 - \frac{1}{2}, \dots, i_d - \frac{1}{2})_{\substack{i_j = \frac{1}{2}, \dots, m_j + \frac{1}{2} \\ i_k = 1, \dots, m_k, k \neq j}}, \\ X^{e,j} &= h(i_1 - \frac{1}{2}, \dots, i_d - \frac{1}{2})_{\substack{i_j = 1, \dots, m_j, \\ i_k = \frac{1}{2}, \dots, m_k + \frac{1}{2}, k \neq j}}, \end{aligned}$$

where $j = 1, \dots, d$. The nodal grid is numbered with half integers, the cell centered grid with integers, the j th face (edge) staggered grid with integers (half integers) except for the j th direction for which half integers (integers) are used.

We denote the discrete analog of the continuous vector field $\mathbf{u} = (u^1, \dots, u^d)$ by $U = (U^1, \dots, U^d)$ where U^k is a grid function approximated on the face-staggered grid $X^{f,k}$. Thus, each of the u^k 's is approximated at different locations which are staggered. We approximate the derivatives $\partial_j u^k$ by

$$\partial_j u^k \approx D_M^{j,k}[U^k].$$

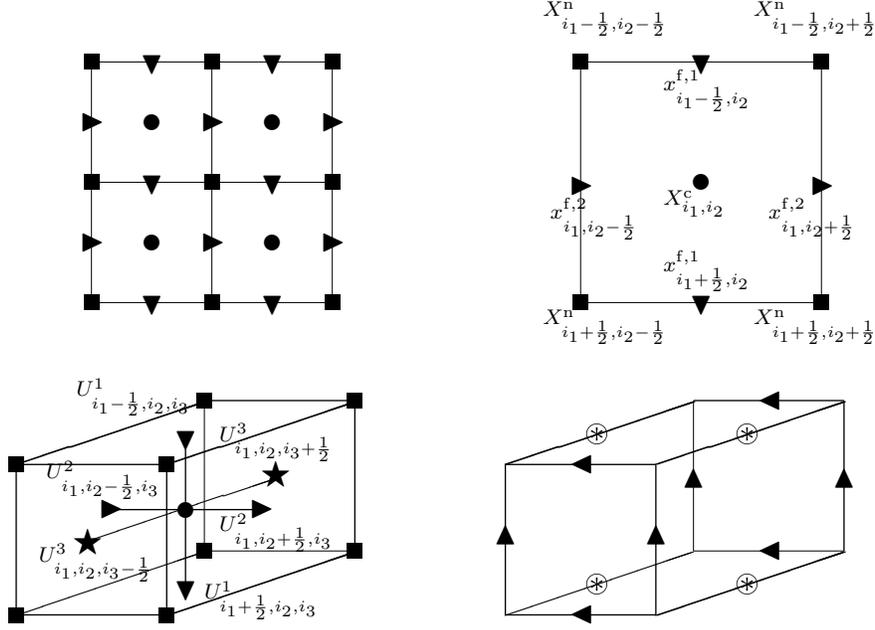


Figure 3: Staggered grids, nodal \blacksquare , cell centered \bullet , face staggered grids (\blacktriangledown in x_1 -, \blacktriangleright in x_2 -, and \blackstar in x_3 -direction), and edge staggered grids (\blacktriangle in x_1 -, \blacktriangleleft in x_2 -, and \otimes in x_3 -direction). TOP LEFT: $d = 2$, four pixels, TOP RIGHT: $d = 2$ pixel (i_1, i_2) with grids, BOTTOM LEFT: $d = 3$, voxel (i_1, i_2, i_3) with face staggered grids and positions of U^1, U^2, U^3 , BOTTOM RIGHT: $d = 3$, edge staggered grids.

Here, the $D_M^{k,k}$'s approximate derivatives in normal direction and the $D_M^{j,k}$'s, $j \neq k$, derivatives in tangential directions. Neglecting obvious indices, the operators are defined by

$$\begin{aligned}
 D_M^{k,k}[U^k]_{\dots, i_k, \dots} &:= \frac{1}{h} (U^k_{\dots, i_k + \frac{1}{2}, \dots} - U^k_{\dots, i_k - \frac{1}{2}, \dots}), \\
 D_M^{j,k}[U^k]_{\dots, i_j + \frac{1}{2}, \dots, i_k + \frac{1}{2}, \dots} &:= \\
 &\frac{1}{h_j} \begin{cases} 0 & i_j = 0, m_j, \\ U^k_{\dots, i_j + 1, \dots, i_k + \frac{1}{2}, \dots} - U^k_{\dots, i_j, \dots, i_k + \frac{1}{2}, \dots} & 0 < i_j < m_j \end{cases} .
 \end{aligned}$$

Hence, $D_M^{k,k}[U^k]$ is located on the cell centered grid but $D_M^{j,k}[U^k]$ is located on the nodal grid for $d = 2$ and on an edge staggered grid for $d = 3$. Note that

no boundary conditions are needed to calculate derivatives in the normal directions and that we have assumed Neumann boundary conditions in the tangential directions.

Remark 3 *To calculate our constraints we will need to approximate sums of products of the form $(\partial_j u_k)(\partial_k u_j)$, see Remark 2. For $d = 2$ the normal derivatives are naturally approximated on the nodal grid. However for $d = 3$ these derivatives are located on edge-staggered grids and therefore are not centered at the same places. For this reason we introduce the averaging operators from the edge-staggered grids to the nodal grid. Ignoring obvious indices, for $\ell = 1, 2, 3$, we have*

$$P_{M,\ell}^{e \rightarrow n}[V]_{i_1+\frac{1}{2}, i_2+\frac{1}{2}, i_3+\frac{1}{2}} := \begin{cases} 0 & i_\ell = 0, m_\ell \\ \frac{1}{2}(V_{\dots, i_\ell, \dots} + V_{\dots, i_\ell+1, \dots}) & 0 < i_\ell < m_\ell \end{cases}.$$

Thus, for any choice of three numbers with $\{j, k, \ell\} = \{1, 2, 3\}$, the projections $P_{M,\ell}^{e \rightarrow n}[D^{j,k}[U^k]]$ are positioned on the nodal grid.

Finally, we introduce an averaging operator $P_M^{n \rightarrow c}$ from the nodal grid to the cell centered grid,

$$P_M^{n \rightarrow c}[V]_{i_1, \dots, i_d} := 2^{-d} \sum_{j_1, \dots, j_d = \pm \frac{1}{2}} V_{i_1+j_1, \dots, i_d+j_d}.$$

For later convenience, we denote by \vec{X} the vector assembled from the lexicographically ordered entries of the array X . Let $\vec{U} := (U^1{}^\top, \dots, U^d{}^\top)^\top$,

$$D^{k,\ell} \vec{U}^\ell := (D_M^{k,\ell}[U^\ell])^\top \quad \text{and} \quad P^{x \rightarrow y} \vec{V} := (P_M^{x \rightarrow y}[V])^\top.$$

Remark 4 *Although the above introduction gives some insight into the actions of the operators $D^{j,k}$ and $P^{x \rightarrow y}$, in particular for coding reasons it might be advantageous to have a compact formal description. Let \otimes denote the Kronecker product of matrices and I_k an identity matrix of appropriate size, then*

$$\begin{aligned} D^{j,k} &= I_1 \otimes \cdots \otimes I_{j-1} \otimes D_V^{j,k} \otimes I_{j+1} \otimes \cdots \otimes I_d, \\ P_\ell^{e \rightarrow n} &= I_1 \otimes \cdots \otimes I_{\ell-1} \otimes P_{V,\ell}^{e \rightarrow n} \otimes I_{\ell+1} \otimes \cdots \otimes I_d, \\ P_\ell^{f \rightarrow c} &= I_1 \otimes \cdots \otimes I_{\ell-1} \otimes P_{V,\ell}^{f \rightarrow c} \otimes I_{\ell+1} \otimes \cdots \otimes I_d, \\ P^{n \rightarrow c} &= P_{V,d}^{f \rightarrow c} \otimes \cdots \otimes P_{V,1}^{f \rightarrow c}, \end{aligned}$$

where the matrices $D_V^{k,k} \in \mathbb{R}^{m_k, m_k+1}$, $D_V^{j,k} \in \mathbb{R}^{m_j+1, m_j}$, $P_{V,\ell}^{e \rightarrow n} \in \mathbb{R}^{m_\ell+1, m_\ell}$, and $P_{V,\ell}^{f \rightarrow c} \in \mathbb{R}^{m_\ell, m_\ell+1}$ are given by

$$D_V^{k,k} := \frac{1}{h} \begin{pmatrix} -1 & 1 & & \\ & \ddots & \ddots & \\ & & -1 & 1 \end{pmatrix}, \quad D_V^{j,k} := \frac{1}{h} \begin{pmatrix} 0 & \cdots & \cdots & 0 \\ -1 & 1 & & \\ & \ddots & \ddots & \\ 0 & \cdots & \cdots & 0 \end{pmatrix},$$

$$P_{V,\ell}^{e \rightarrow n} := \frac{1}{2} \begin{pmatrix} 0 & \cdots & \cdots & 0 \\ 1 & 1 & & \\ & \ddots & \ddots & \\ & & 1 & 1 \\ 0 & \cdots & \cdots & 0 \end{pmatrix}, \quad P_{V,\ell}^{f \rightarrow c} := \frac{1}{2} \begin{pmatrix} 1 & 1 & & \\ & \ddots & \ddots & \\ & & 1 & 1 \end{pmatrix}.$$

3.2 Discretization of the building blocks \mathcal{D} , \mathcal{S} , and \mathcal{C}

For the particular building blocks we derive discrete analogs. Let $\vec{X}^c = (\vec{X}_1^{c\top}, \dots, \vec{X}_d^{c\top})^\top$, $\vec{R} = R(\vec{X}^c)$, and

$$\vec{T}(\vec{U}) = T(\vec{X}_1^c + P_1^{f \rightarrow c} \vec{U}^1, \dots, \vec{X}_d^c + P_d^{f \rightarrow c} \vec{U}^d).$$

Note that $\vec{T}(\vec{U})$ is the discrete analog of the image $T(\mathbf{x} + \mathbf{u}(\mathbf{x}))$ as a function of \mathbf{u} . Since T is assumed to be a smooth function, $T(\mathbf{x})$ can be evaluated for any \mathbf{x} . In our implementation we use a B-spline interpolation scheme. The assumption on T to be differentiable is for ease of presentation only. In the continuous formulation (3), derivatives appear only in a weak formulation. Thus, all we need is the existence of a distributional derivative of T .

We denote the Jacobian of \vec{T} by

$$J(\vec{U}) := \frac{\partial \vec{T}}{\partial \vec{U}}(\vec{U}) = \left(\text{diag}((P_1^{f \rightarrow c})^\top \partial_1 T), \dots, \text{diag}((P_d^{f \rightarrow c})^\top \partial_d T) \right), \quad (15)$$

where the partial derivatives $\partial_j T$ have to be evaluated at $(\vec{Y}_1, \dots, \vec{Y}_d)$, $\vec{Y}_j = \vec{X}_j^c + Q_j \vec{U}_j$.

Discretizing \mathcal{D} The discretization of $\mathcal{D}[R, T; \mathbf{u}]$ (2) is straightforward,

$$\mathbf{D}(\vec{U}) := \frac{1}{2} \|\vec{T}(\vec{U}) - \vec{R}\|_2^2$$

and its derivative, which is also known and interpreted as a force field \vec{F} (see, e.g. [26]), is

$$\vec{F}(\vec{U}) := \mathbf{D}_{\vec{U}}(\vec{U}) = J(\vec{U})^\top (\vec{T}(\vec{U}) - \vec{R}). \quad (16)$$

Discretizing \mathcal{S} Following (7), all we need are discretizations $\nabla^h \times$ and $\nabla^h \cdot$ of $\nabla \times$ and $\nabla \cdot$, respectively. Since the discretization for $\nabla \times$ and $\nabla \cdot$ is composed from first order derivatives, for $d = 2, 3$, we obtain

$$\begin{pmatrix} \nabla_2^h \times \\ \nabla_2^h \cdot \end{pmatrix} := \begin{pmatrix} D^{2,1} & -D^{1,2} \\ D^{1,1} & D^{2,2} \end{pmatrix} \text{ and } \begin{pmatrix} \nabla_3^h \times \\ \nabla_3^h \cdot \end{pmatrix} := \begin{pmatrix} 0 & -D^{3,1} & D^{2,3} \\ D^{3,1} & 0 & -D^{1,3} \\ -D^{2,1} & D^{1,2} & 0 \\ D^{1,1} & D^{2,2} & D^{3,3} \end{pmatrix}.$$

With $A = \mathbf{B}^\top \mathbf{B}$, $\mathbf{B} = \text{diag}(\sqrt{\mu}, \sqrt{2\mu + \lambda})(\nabla^h \times, \nabla^h \cdot)^\top$, the discretization of $\mathcal{S}[\mathbf{u}]$ (5) is given by

$$\mathcal{S}(\vec{U}) := \frac{1}{2} \vec{U}^\top A \vec{U}, \quad \mathcal{S}_{\vec{U}}(\vec{U}) = A \vec{U}.$$

Here, we imposed zero Neumann boundary conditions in tangential directions.

Discretizing \mathcal{C} In our discretization of the volume preserving constraints we exploit the weak form of (8) applied to an elementary volume $V \subset \mathbb{R}^d$, which is a pixel for $d = 2$ and a voxel for $d = 3$. Therefore, the vertices belong to the nodal grid. Fig. 4 displays a deformed volume $\varphi(V)$, where $\varphi(\mathbf{x}) = \mathbf{x} + \mathbf{u}(\mathbf{x})$.

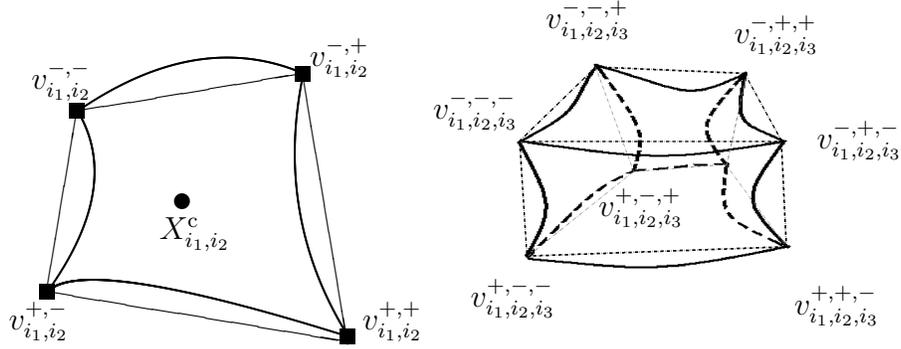


Figure 4: Deformed elementary volume and its approximation, LEFT: a pixel for $d = 2$, RIGHT: a voxel for $d = 3$.

We approximate the volume of $\varphi(V)$ by the volume of the set \tilde{V} spanned by the linearly interpolated $d - 1$ dimensional surfaces, where the interpolation points are vertices. For $d = 2$, the set \tilde{V} is bounded by four straight

lines, and for $d = 3$ it is bounded by twelve triangles. Assuming that the second derivatives of φ are bounded, the interpolation error and hence the approximation error is of order h^2 .

For $d = 2$, we end up with

$$\int_{\varphi(V_{i_1, i_2})} dV \mathbf{x} \approx \int_{\tilde{V}_{i_1, i_2}} d\mathbf{x} = \frac{1}{2} \det(v_{i_1, i_2}^{+,+} - v_{i_1, i_2}^{-,-}, v_{i_1, i_2}^{-,+} - v_{i_1, i_2}^{+,-}),$$

where the determinant is computed for the matrix spanned by the two vectors, the two vectors are the differences between the four vertices

$$v_{i_1, i_2}^{\pm, \pm} = X_{i_1 \pm \frac{1}{2}, i_2 \pm \frac{1}{2}}^n + (U_{i_1 \pm \frac{1}{2}, i_2 \pm \frac{1}{2}}^1, U_{i_1 \pm \frac{1}{2}, i_2 \pm \frac{1}{2}}^2)^\top,$$

and $U_{i_1 \pm \frac{1}{2}, i_2 \pm \frac{1}{2}}^j$ is averaged over the staggered grid. To be precise, we use two average operators $P_j^{f \rightarrow n}$ to map the U^j 's from the face staggered to the nodal grid and four projections $E_{m_j}^\pm$ to map to the top left $(-, -)$, bottom left $(+, -)$, top right $(-, +)$, and bottom right positions $(+, +)$ relative to the cell center. Using vectorized quantities, we have $\vec{U}^{j, \pm, \pm} = P_j^{\pm, \pm} \vec{U}^j$, or, precisely,

$$\begin{aligned} \vec{U}^{1,+,+} &= (E_2^+ P_2^{f \rightarrow n}) \otimes E_1^+ \cdot \vec{U}^1, & \vec{U}^{1,+,-} &= (E_2^- P_2^{f \rightarrow n}) \otimes E_1^+ \cdot \vec{U}^1, \\ \vec{U}^{1,-,+} &= (E_2^+ P_2^{f \rightarrow n}) \otimes E_1^- \cdot \vec{U}^1, & \vec{U}^{1,-,-} &= (E_2^- P_2^{f \rightarrow n}) \otimes E_1^- \cdot \vec{U}^1, \\ \vec{U}^{2,+,+} &= E_2^+ \otimes (E_1^+ P_1^{f \rightarrow n}) \cdot \vec{U}^2, & \vec{U}^{2,+,-} &= E_2^- \otimes (E_1^+ P_1^{f \rightarrow n}) \cdot \vec{U}^2, \\ \vec{U}^{2,-,+} &= E_2^+ \otimes (E_1^- P_1^{f \rightarrow n}) \cdot \vec{U}^2, & \vec{U}^{2,-,-} &= E_2^- \otimes (E_1^- P_1^{f \rightarrow n}) \cdot \vec{U}^2, \end{aligned}$$

where $E_j^- = I_{m_j+1}(1:m_j, 1:m_j+1)$, $E_j^+ = I_{m_j+1}(2:m_j+1, 1:m_j+1) \in \mathbb{R}^{m_j, m_j+1}$, and

$$P_j^{f \rightarrow n} = \frac{1}{2} \begin{pmatrix} 2 & & & & \\ 1 & 1 & & & \\ & \ddots & \ddots & & \\ & & & 1 & 1 \\ & & & & 2 \end{pmatrix} \in \mathbb{R}^{m_j+1, m_j}.$$

Using these projections and omitting the indices, we have

$$\int_{\tilde{V}} dV = \frac{1}{2} \det \begin{pmatrix} h + \vec{U}^{1,++} - \vec{U}^{1,--} & \vec{U}^{1,-+} - h - \vec{U}^{1,+ -} \\ h + \vec{U}^{2,++} - \vec{U}^{2,--} & h + \vec{U}^{2,-+} - \vec{U}^{2,+ -} \end{pmatrix},$$

and hence

$$\frac{1}{h^2} \int_{\tilde{V}} dV - 1 = \frac{1}{2h} \left(\vec{U}^{1,++} - \vec{U}^{1,--} + \vec{U}^{1,+} - \vec{U}^{1,-} \right) \quad (17)$$

$$+ \frac{1}{2h} \left(\vec{U}^{2,++} - \vec{U}^{2,--} + \vec{U}^{2,+} - \vec{U}^{2,-} \right) \quad (18)$$

$$+ \frac{1}{2} \left((\vec{U}^{1,++} - \vec{U}^{1,--})(\vec{U}^{2,+} - \vec{U}^{2,-}) \right. \\ \left. - (\vec{U}^{1,+} - \vec{U}^{1,-})(\vec{U}^{2,++} - \vec{U}^{2,--}) \right) \quad (19)$$

and with $c_{i_1, i_2}(\vec{U}) = \frac{1}{h^2} \int_{\tilde{V}_{i_1, i_2}} dV - 1$, the discrete constraints are defined by

$$\mathbf{C}(\vec{U}) = (c_{i_1, i_2}(\vec{U}))^7. \quad (20)$$

Remark 5 *The terms in (17) and (18) are approximations to the derivatives $\partial_1 u^1$ and $\partial_2 u^2$ on the cell centered grids using the stencils $S^{1,1}$ and $S^{2,2}$ for U^1 and U^2 on the face staggered grids, respectively. Moreover, (19) is an approximation to $\det(\nabla \mathbf{u})$. With $a := U^{+, -} - U^{-, -}$, $b := U^{-, +} - U^{-, -}$, $c := U^{+, +} - U^{-, +}$, $d := U^{+, +} - U^{+, -}$ we find*

$$\frac{1}{2} \det(U^{+, +} - U^{-, -}, U^{-, +} - U^{+, -}) \\ = \frac{1}{2} \det(b + c, b - a) = \det((a + c)/2, (b + d)/2),$$

where $(a + c)/2$ is an approximation to $\partial_1 \mathbf{u}$ and $(b + d)/2$ is an approximation to $\partial_2 \mathbf{u}$, respectively. The stencils are

$$S^{1,1} = \frac{1}{4} \begin{pmatrix} -1 & -2 & -1 \\ 1 & 2 & 1 \end{pmatrix}, \quad S^{1,2} = \frac{1}{4} \begin{pmatrix} -1 & -1 \\ 0 & 0 \\ 1 & 1 \end{pmatrix}, \\ S^{2,1} = \frac{1}{4} \begin{pmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix}, \quad S^{2,2} = \frac{1}{4} \begin{pmatrix} -1 & 1 \\ -2 & 2 \\ -1 & 1 \end{pmatrix}.$$

These stencils are related to Sobel operators; cf., e.g. [16]. However, based on the staggered grids, for the normal derivatives short differences are used. The stencil in normal directions applied to a vector field is a finite element approximation of the divergence,

$$\nabla^h \cdot (U^1, U^2) \approx S^{1,1} U^1 + S^{2,2} U^2.$$

The long stencils $S^{1,2}$ and $S^{2,1}$ approximate the derivatives in the tangential direction. Therefore, our discretization can be thought of as a finite difference or a finite element discretization of the continuous quantity VP (9).

The derivative $\mathbf{C}_{\vec{v}}$ is computed directly:

$$\begin{aligned}
\mathbf{C}_{\vec{v}} &= \begin{pmatrix} P_1^{++} - P_1^{-,-} + P_1^{+,-} - P_1^{-,+} \\ P_2^{++} - P_2^{-,-} + P_2^{+,-} - P_2^{-,+} \end{pmatrix} \\
&+ \begin{pmatrix} \text{diag}((P_2^{-,+} - P_2^{+,-}) \cdot \vec{U}^2) \cdot (P_1^{+,+} - P_1^{-,-}) \\ \text{diag}((P_1^{+,+} - P_1^{-,-}) \cdot \vec{U}^1) \cdot (P_2^{-,+} - P_2^{+,-}) \end{pmatrix} \\
&+ \begin{pmatrix} \text{diag}((P_2^{+,+} - P_2^{-,-}) \cdot \vec{U}^2) \cdot (P_1^{-,+} - P_1^{+,-}) \\ \text{diag}((P_1^{-,+} - P_1^{+,-}) \cdot \vec{U}^1) \cdot (P_2^{+,+} - P_2^{-,-}) \end{pmatrix}. \quad (21)
\end{aligned}$$

For $d = 3$, the volume is computed by summing the volume of six pyramids with bases top $(v^{-,\pm,\pm})$, bottom $(v^{+,\pm,\pm})$, left $(v^{\pm,-,\pm})$, right $(v^{\pm,+,\pm})$, front $(v^{\pm,\pm,-})$, and back $(v^{\pm,\pm,+})$, and tip $t := X_{i_1, i_2, i_3}^c + U_{i_1, i_2, i_3}$. The eight vertices are

$$v_{i_1, i_2, i_3}^{\pm, \pm, \pm} = X_{i_1 \pm \frac{1}{2}, i_2 \pm \frac{1}{2}, i_3 \pm \frac{1}{2}}^n + (U_{i_1 \pm \frac{1}{2}, i_2 \pm \frac{1}{2}, i_3 \pm \frac{1}{2}}^1, \dots, U_{i_1 \pm \frac{1}{2}, i_2 \pm \frac{1}{2}, i_3 \pm \frac{1}{2}}^3)^\top,$$

and $U_{i_1 \pm \frac{1}{2}, i_2 \pm \frac{1}{2}, i_3 \pm \frac{1}{2}}^j$ is averaged over the staggered grid in an analogous way as described for $d = 2$.

Each pyramid is a conglomeration of two tetrahedra. The volume $T(a, b, c, t)$ of a tetrahedron with vertices a, b, c , and t is

$$T(a, b, c, t) = \frac{1}{6} \det(b - a, c - a, t - a),$$

where the vertices have to be numbered such that the determinant is non-negative. Therefore,

$$\begin{aligned}
&6(T(v^{-,+}, v^{-,-}, v^{+,-}, t) + T(v^{-,+}, v^{+,-}, v^{+,+}, t)) \\
&= \det(v^{-,-} - v^{-,+}, v^{+,-} - v^{-,+}, t - v^{-,+}) \\
&\quad + \det(v^{+,-} - v^{-,+}, v^{+,+} - v^{-,+}, t - v^{-,+}) \\
&= \det(v^{+,+} - v^{-,-}, v^{-,+} - v^{+,-}, t - v^{-,+})
\end{aligned}$$

and thus

$$\begin{aligned}
6\tilde{V} &= \det(v^{-,+,+} - v^{-,-,-}, v^{-,-,+} - v^{-,+,+}, t - v^{-,-,+}) \\
&+ \det(v^{+,+,+} - v^{+,-,-}, v^{+,-,+} - v^{+,+,+}, t - v^{+,-,+}) \\
&+ \det(v^{+,-,+} - v^{-,-,-}, v^{-,-,+} - v^{+,-,+}, t - v^{-,-,+}) \\
&+ \det(v^{+,+,+} - v^{-,+,+}, v^{-,+,+} - v^{+,+,+}, t - v^{-,+,+}) \\
&+ \det(v^{+,-,+} - v^{-,-,-}, v^{-,+,+} - v^{+,-,+}, t - v^{-,+,+}) \\
&+ \det(v^{+,+,+} - v^{-,-,+}, v^{-,-,+} - v^{+,-,+}, t - v^{-,+,+}).
\end{aligned}$$

With $c_{i_1, i_2, i_3} := \frac{1}{h^3} \tilde{V}_{i_1, i_2, i_3} - 1$, the discrete constraints are defined analogously to (20). The projections $\vec{U}^{j, \pm, \pm, \pm} = P_j^{\pm, \pm, \pm} \vec{U}^j$ are similar to the ones used for $d = 2$, therefore we give just one example,

$$\vec{U}^{1, +, +, +} = (E_3^+ P_3^{f \rightarrow n}) \otimes (E_2^+ P_2^{f \rightarrow n}) \otimes E_1^+ \cdot \vec{U}^1$$

Since the complete formula for the three-dimensional case is lengthy but its derivation is along the same line as the one for two dimensions, we omit the details. However, it is important to note that for $d = 2$, the resulting nonlinear equations are in general quadratic, but for $d = 3$ they are cubic.

4 Solving the discrete optimization problem

We are now ready to phrase the discrete analog of the image registration problem (1),

$$\mathbf{J}(\vec{U}) := \mathbf{D}(\vec{U}) + \alpha \mathbf{S}(\vec{U}) = \min \quad (22a)$$

$$\text{subject to } \mathbf{C}(\vec{U}) = 0. \quad (22b)$$

In order to solve this problem numerically we use the framework of Sequential Quadratic Programming (SQP); see [27] for a detailed discussion. Let \vec{P} be a cell-centered vector of Lagrange multipliers. The Lagrangian of the problem is

$$\mathbf{L}(\vec{U}, \vec{P}) = \mathbf{D}(\vec{U}) + \frac{\alpha}{2} \vec{U}^\top A \vec{U} + \mathbf{C}(\vec{U})^\top \vec{P}.$$

Differentiating with respect to \vec{U} and \vec{P} , we obtain the following discrete version of the Euler-Lagrange equations (13)

$$0 = \mathbf{L}_{\vec{U}}(\vec{U}, \vec{P}) = \vec{F}(\vec{U}) + \alpha A \vec{U} + \mathbf{C}_{\vec{U}}(\vec{U})^\top \vec{P}, \quad (23a)$$

$$0 = \mathbf{L}_{\vec{P}}(\vec{U}, \vec{P}) = \mathbf{C}(\vec{U}). \quad (23b)$$

We can now solve the nonlinear system (23) numerically by using a Newton-type method. Approximating the (1,1) block of the Hessian by

$$H := \alpha A + J^\top J, \quad (24)$$

where J is defined in (15), we obtain the following linear system of equations to be solved at each iteration:

$$\begin{pmatrix} H & \mathbf{C}_{\vec{U}}^\top \\ \mathbf{C}_{\vec{U}} & 0 \end{pmatrix} \begin{pmatrix} s_{\vec{U}} \\ s_{\vec{P}} \end{pmatrix} = \begin{pmatrix} \mathbf{L}_{\vec{U}}(\vec{U}, \vec{P}) \\ \mathbf{L}_{\vec{P}}(\vec{U}, \vec{P}) \end{pmatrix}. \quad (25)$$

The system (25) is a so-called Karush-Kuhn-Tucker (KKT) system; it is symmetric but indefinite. Solving KKT systems is a well known challenge. Similar systems arise for example in fluid dynamics (cf., e.g., [17, 33]) and the solution for this case has been addressed by many authors; see, e.g., [33, 31, 32]. Nevertheless, the robust and effective solution of such systems is still an open research topic. Here we have used MINRES [28] with a block diagonal preconditioning as proposed in [31]. This preconditioner can be written as

$$M = \begin{pmatrix} H & 0 \\ 0 & \hat{S} \end{pmatrix},$$

where \hat{S} is an approximation to the Schur complement

$$S := \mathbf{C}_{\vec{v}} H^{-1} \mathbf{C}_{\vec{v}}^\top.$$

Here, we use the approximation suggested in [31],

$$\hat{S}^{-1} = (\mathbf{C}_{\vec{v}} \mathbf{C}_{\vec{v}}^\top)^{-1} \mathbf{C}_{\vec{v}} H \mathbf{C}_{\vec{v}}^\top (\mathbf{C}_{\vec{v}} \mathbf{C}_{\vec{v}}^\top)^{-1}. \quad (26)$$

Note that $(\mathbf{C}_{\vec{v}} \mathbf{C}_{\vec{v}}^\top)^{-1} \mathbf{C}_{\vec{v}}$ is the pseudo-inverse of $\mathbf{C}_{\vec{v}}$. The application of the preconditioner only involves a multiplication of H^{-1} and \hat{S}^{-1} with a vector. However, an efficient numerical scheme is not straightforward and will be addressed in a forthcoming paper. Here, we use a multigrid approach.

After the KKT system has been solved, we update \vec{U} by setting

$$\vec{U} \leftarrow \vec{U} + \gamma s_{\vec{v}}.$$

As it is common in SQP algorithms [27], the parameter γ is chosen such that the L_1 merit function

$$\text{merit}_{\text{KKT}}(\vec{U}) := \mathbf{D}(\vec{U}) + \alpha \mathbf{S}(\vec{U}) + \theta \|\mathbf{C}(\vec{U})\|_1 \quad (27)$$

decreases, where $\theta := \|\vec{P}\|_\infty + \theta_{\min}$ with a fixed parameter θ_{\min} , $\|x\|_1 = \sum_j |x_j|$, and \vec{P} is the least squares multiplier computed by solving

$$(\mathbf{C}_{\vec{v}} \mathbf{C}_{\vec{v}}^\top) \vec{P} = \mathbf{C}_{\vec{v}} (\mathbf{D}_{\vec{v}} + \alpha \mathbf{S}_{\vec{v}}). \quad (28)$$

In order to avoid to be too far from feasibility we project the intermediate \vec{U} to the constraints. Here we take advantage of a second merit function

$$\text{merit}_{\text{C}}(\vec{U}) := \|\mathbf{C}(\vec{U})\|_2^2. \quad (29)$$

If $\text{merit}_C(\vec{U}) \geq \text{tol}_C$, we calculate a step \vec{U}_P such that ideally

$$\mathbf{C}(\vec{U} + \vec{U}_P) \approx \mathbf{C}(\vec{U}) + \mathbf{C}_{\vec{U}}(\vec{U})\vec{U}_P = 0.$$

If $\mathbf{C}_{\vec{U}}$ has full rank a solution is given by $\vec{U}_P = \mathbf{C}_{\vec{U}}(\vec{U})^\top \vec{W}$, where \vec{W} is the solution of the projection system

$$(\mathbf{C}_{\vec{U}} \mathbf{C}_{\vec{U}}^\top) \cdot \vec{W} = -\mathbf{C}(\vec{U}). \quad (30)$$

Note that the very same system has already be addressed in the Schur complement approximation (26) and (28).

The above step is repeated until convergence of the optimization process, which is measured by the change of \vec{U} . In order to find an appropriate regularization parameter, we solve (22) for a few values of α , where we start with a large α and slowly decrease it until our stopping criterion is fulfilled. For the results presented in Section 5, the stopping criterion is based on a visual inspection of the images. Our numerical scheme is summarized in Alg. 1.

5 Numerical examples

5.1 The Blob

To illustrate the potential of the volume preserving registration we present a synthetic example; see Fig. 5. The reference image (top right) shows an elliptic global structure which contains a small almost circular object. The template (top left) shows a rotated version of the global ellipse, where the inner structure is translated and considerably enlarged. Note that this example mimics the situation for contrast enhanced images: slightly deformed global structures, where inner structures may change drastically due to contrast uptake.

As it is apparent from Fig. 5, the unconstrained registration gives very good results if we are looking at the difference between the reference and deformed template images alone. However, as expected, the inner structure has been reduced so as to fit the one in the reference image. This results in a drastic change of volume, which can be observe from the visualization of a part of the grid in Fig. 5 (middle right) corresponding to a region of interest emphasized in the template image (top left). Thus, for contrast enhanced

Algorithm 1 Volume Preserving Image Registration: $\vec{U} \leftarrow \text{VPIR}(R, T)$.

- 1: Set $k \leftarrow 0$, $\vec{U} \leftarrow 0$, $\vec{P} \leftarrow 0$, $n \leftarrow \text{length}(\vec{U})$.
- 2: **for** $k = 0, \dots$ **do**
- 3: Compute \mathbf{D} , J , $\mathbf{D}_{\vec{U}}$, \mathbf{S} , $\mathbf{S}_{\vec{U}}$, \mathbf{C} , and $\mathbf{C}_{\vec{U}}$, set $\vec{U}_{\text{old}} = \vec{U}$.
- 4: Set $\begin{pmatrix} L_{\vec{U}} \\ L_{\vec{P}} \end{pmatrix} \leftarrow \begin{pmatrix} \mathbf{D}_{\vec{U}} + \alpha \mathbf{S}_{\vec{U}} + \mathbf{C}_{\vec{U}}^\top \vec{P} \\ \mathbf{C} \end{pmatrix}$.
 { Computation of the SQP step }
- 5: Solve the KKT system (25) for $s_{\vec{U}}$ and $s_{\vec{P}}$
- 6: Solve for the Lagrange-multiplier \vec{P} , cf. (28); set $\theta \leftarrow \|\vec{P}\|_\infty$
- 7: $\vec{U} \leftarrow \text{LS}(\vec{U}, s_{\vec{U}}, \text{merit}_{\text{KKT}}, L_{\vec{U}})$; see Alg. 2.
- 8: Update \mathbf{C} and $\mathbf{C}_{\vec{U}}$.
 {Computation of the projection onto the constraints}
- 9: **while** $\text{merit}_{\mathbf{C}} > \text{tol}_{\mathbf{C}}$ **do**
- 10: solve $(\mathbf{C}_{\vec{U}} \ \mathbf{C}_{\vec{U}}^\top) \vec{W} = -\mathbf{C}$, set $\vec{U}_P \leftarrow \mathbf{C}_{\vec{U}}^\top \vec{W}$.
- 11: $\vec{U} \leftarrow \text{LS}(\vec{U}, \vec{U}_P, \text{merit}_{\mathbf{C}}, \mathbf{C}_{\vec{U}}^\top \mathbf{C})$; see Alg. 2.
- 12: **if** $\|\vec{U}_{\text{old}} - \vec{U}\| \leq \text{tol}_U$ **then**
- 13: return \vec{U} , done.
- 14: **end if**
- 15: **end while**
- 16: **end for**

Algorithm 2 Armijo's line search: $\vec{U} \leftarrow \text{LS}(\vec{U}, s_{\vec{U}}, \text{merit}, \vec{G})$

Set $j \leftarrow 0$, $\gamma \leftarrow 1$, and $\eta \leftarrow 10^{-5}$.

while true **do**

 Set $\vec{U}_t \leftarrow \vec{U} + \gamma s_{\vec{U}}$.

if $\text{merit}(\vec{U}_t) < \text{merit}(\vec{U}) + \eta \gamma \vec{G}^\top s_{\vec{U}}$ **then**

 break

 {step reduces merit function}

end if

if $j > \max_\ell$ **then**

 error

 {step not successful after \max_ℓ line search steps}

end if

 Set $\gamma \leftarrow \gamma/2$, and $j \leftarrow j + 1$.

end while

Set $\vec{U} \leftarrow \vec{U}_t$.

images, the registration gives meaningless results, though the difference is small.

Fig. 5 also shows the results of the volume preserving registration (bottom left). As is apparent from this figure, the global deformation has been resolved, the inner ellipse has been moved to match the inner ellipse in the reference image. However, the volume of the inner ellipse has not been altered, which leads to a larger difference as in the unconstrained case but also to a more realistic registration; see also the deformed grid (bottom right).

In order to compare these results, we choose $\alpha = 10^3$ and stop after convergence, which occurs after at most 25 iterations for both registrations. Note that $h = 1$ in our implementation. The values for the difference \mathbf{D} and the constraints \mathbf{C} for the un- and VP-constrained registration are summarized in Table 1.

5.2 MRI scans

In our second example, we discuss results obtained for the images shown in Fig. 1. Fig. 6 shows the results after two (2nd row) and ten iterations (3rd row) of the unconstrained registration as well as after ten iterations of the VP constrained registration (4th row). The numerical results are summarized in Table 1. After ten iteration both schemes have converged.

Table 1: Numerical results for the un- and VP-constrained registrations; k is the number of iteration performed.

		α	k	$\mathbf{D}(\vec{U}^{(k)})/\mathbf{D}(0)$	$\ \mathbf{C}(\vec{U}^{(k)})\ _\infty$
blob	unconstrained	10^3	25	0.21	0.87
	VP constrained	10^3	25	0.73	$\leq 10^{-6}$
MRI	unconstrained	10^5	2	0.81	1.36
	unconstrained	10^5	10	0.78	1.36
	VP constrained	10^5	10	0.87	$\leq 10^{-6}$

Although the numbers (cf. Table 1) indicate a larger reduction of the difference by the unconstrained registrations, the ranking is not so clear if one looks at the difference images, cf. Fig. 6. Here, the difference after ten steps un- and VP constrained registration looks pretty much the same. After two steps of the unconstrained registration the bright spot in the top part of the image has not been resolved satisfiably. The explanation is that small

spots which are related to noise in the MRI images and hardly visible in the images are registered in the unconstrained registration. This leads to a large reduction though it is hardly visible. To remove this small spots, the volume has to be changed locally. However, the registration of these small spots does not contribute to a meaningful solution for this problem.

6 Summary

In this paper we have presented numerical methods for the solution of volume preserving image registration problems. We have developed a stable discretization to the optimization problem and used a variant of Sequential Quadratic Programming to solve the problem. This results in a highly effective algorithm for the solution of the problem. We have tested our algorithm on real medical data as well as on synthetic data.

Our formulation opens up a few avenues of research such as effective solvers for the KKT systems and inexact SQP methods. These issues will be addressed in a subsequent paper.

Acknowledgements: Jan Modersitzki was supported by the US National Institutes of Health under Grant NIHR01 HL 068904. We are indebted to Michele Benzi for suggestions and discussions.

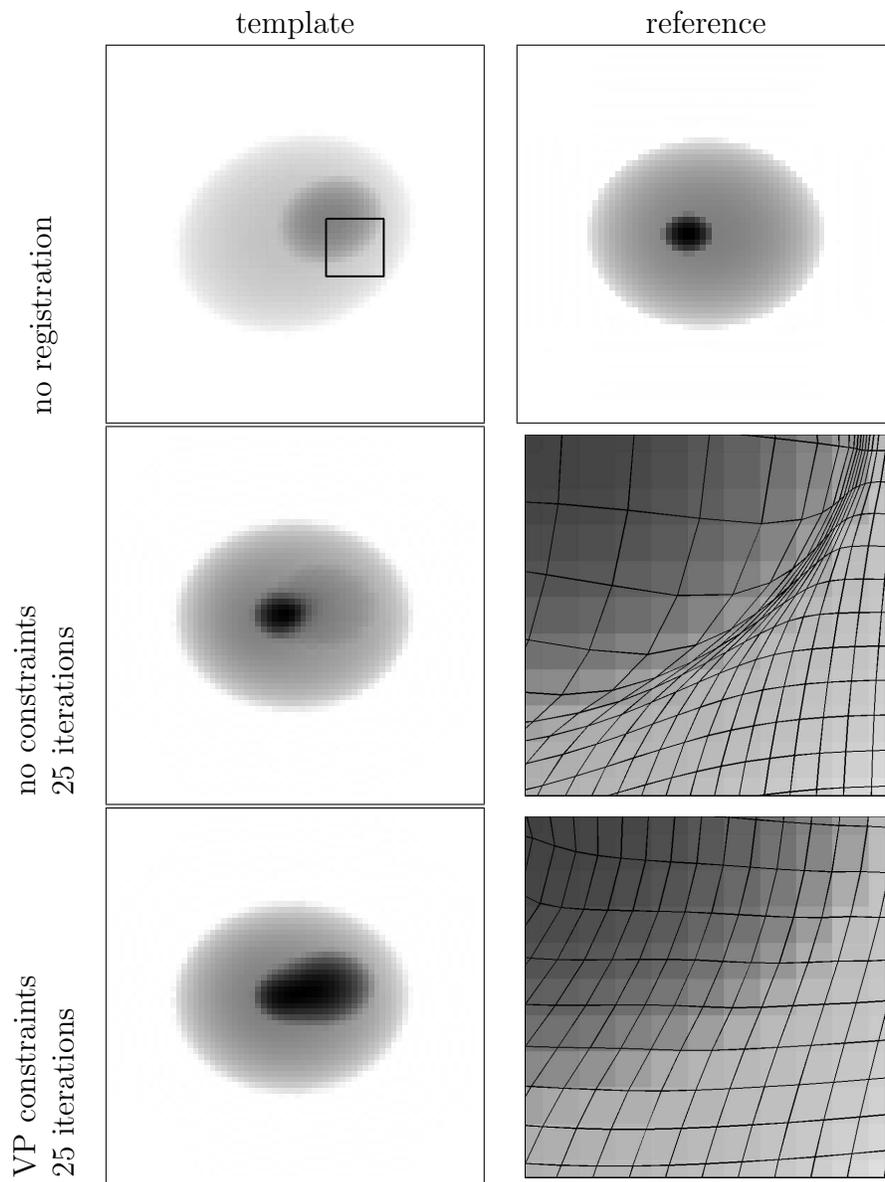


Figure 5: Synthetic example, LEFT COLUMN: deformed template, RIGHT COLUMN: reference and detail of the deformed grid; TOP ROW: template and reference, no registration, MIDDLE ROW: deformed template and details with grid after unconstrained registration, BOTTOM ROW: deformed template and details with grid after VP constrained registration. For both schemes, we choose $\alpha = 10^3$ and stopped after 25 iterations.

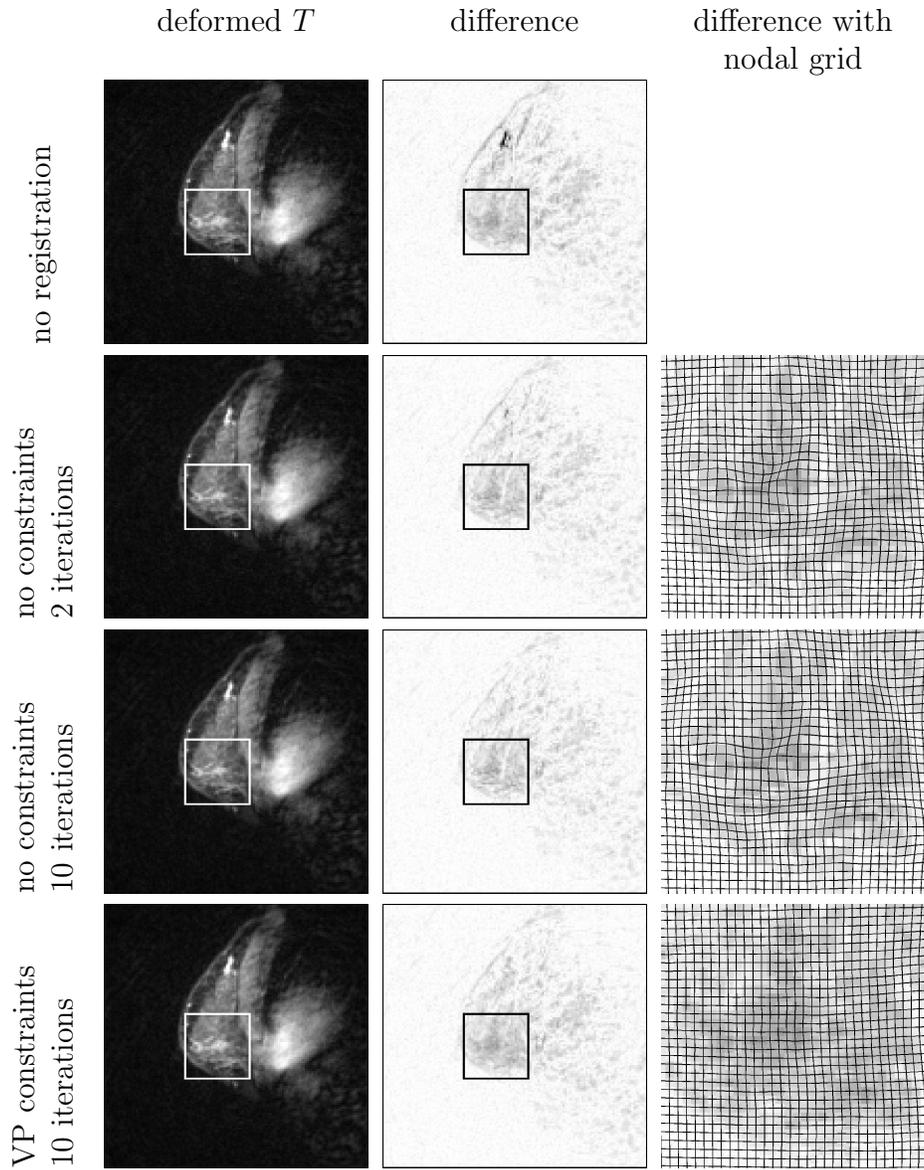


Figure 6: Registration results for the images of Fig. 1. LEFT COLUMN deformed template images T_u , MIDDLE COLUMN difference image $|R - T_u|$ with region of interest (ROI), RIGHT COLUMN ROI with nodal grid, vertices connected by straight lines ; ROW 1: no registration, ROW 2: no constraints two iterations, ROW 3: no constraints ten iterations, and ROW 4: volume preserving constraints ten iterations.

References

- [1] Fred L. Bookstein, *Principal warps: Thin-plate splines and the decomposition of deformations*, IEEE Transactions on Pattern Analysis and Machine Intelligence **11** (1989), no. 6, 567–585.
- [2] F. Brezzi and M. Fortin, *Mixed and hybrid finite element methods*, Springer, 1991.
- [3] Morten Bro-Nielsen, *Medical image registration and surgery simulation*, Ph.D. thesis, IMM, Technical University of Denmark, 1996.
- [4] Chaim Broit, *Optimal registration of deformed images*, Ph.D. thesis, Computer and Information Science, University of Pennsylvania, 1981.
- [5] Gary E. Christensen and H. J. Johnson, *Consistent image registration*, IEEE Transaction on Medical Imaging **20** (2001), no. 7, 568–582.
- [6] Gary Edward Christensen, *Deformable shape models for anatomy*, Ph.D. thesis, Sever Institute of Technology, Washington University, 1994.
- [7] A. Collignon, A. Vandermeulen, P. Suetens, and G. Marchal, *3d multi-modality medical image registration based on information theory*, Kluwer Academic Publishers: Computational Imaging and Vision **3** (1995), 263–274.
- [8] R. Courant and David Hilbert, *Methods of mathematical physics*, vol. II, Wiley, New York, 1962.
- [9] Bernd Fischer and Jan Modersitzki, *Fast inversion of matrices arising in image processing*, Num. Algo. **22** (1999), 1–11.
- [10] ———, *Fast diffusion registration*, AMS Contemporary Mathematics, Inverse Problems, Image Analysis, and Medical Imaging, vol. 313, 2002, pp. 117–129.
- [11] ———, *Combination of automatic non-rigid and landmark based registration: the best of both worlds*, Medical Imaging 2003: Image Processing (M. Sonka and J.M. Fitzpatrick, eds.), Proceedings of the SPIE 5032, 2003, pp. 1037–1048.

- [12] ———, *Curvature based image registration*, J. of Mathematical Imaging and Vision **18** (2003), no. 1, 81–85.
- [13] J. M. Fitzpatrick, D. L. G. Hill, and C. R. Maurer Jr., *Image registration*, Handbook of Medical Imaging, Volume 2: Medical Image Processing and Analysis (M. Sonka and J. M. Fitzpatrick, eds.), SPIE, 2000, pp. 447–513.
- [14] C. A. J. Fletcher, *Computational techniques for fluid dynamics*, vol. II, Springer, 1988.
- [15] James C. Gee, David R. Haynor, Lionel Le Briquer, and Ruzena Bajcsy, *Advances in elastic matching theory and its implementation*, CVRMed, 1997, pp. 63–72.
- [16] Rafael C. Gonzalez and Richard E. Woods, *Digital image processing*, Addison-Wesley, 1993.
- [17] P. M. Gresho and R. L. Sani, *On pressure boundary conditions for the incompressible Navier-Stokes equations*, Internat. J. Numer. Methods Fluids **7** (1987), 1111–1145.
- [18] E. Haber and U. Ascher, *Fast finite volume simulation of 3D electromagnetic problems with highly discontinuous coefficients*, SIAM J. Scient. Comput. **22** (2001), 1943–1961.
- [19] Steven Haker and Allen Tannenbaum, *Optimal transport and image warping*, MICCAI 2001 (2001), 120–127.
- [20] Stefan Heldmann, Oliver Mahnke, Daniel Potts, Jan Modersitzki, and Bernd Fischer, *Fast computation of mutual information in a variational image registration approach*, Bildverarbeitung für die Medizin 2004, Springer, 2004, Accepted for publication, pp. 1–5.
- [21] Stefan Henn and Kristian Witsch, *Iterative multigrid regularization techniques for image matching*, SIAM J. on Scientific Comp. **23** (2001), no. 4, 1077–1093.
- [22] G. Hermosillo, *Variational methods for multimodal image matching*, Ph.D. thesis, Université de Nice, France, 2002.

- [23] B. K. P. Horn and B. G. Schunck, *Determining optical flow*, Artificial Intelligence **17** (1981), 185–204.
- [24] J.M. Hyman and M. Shashkov, *The orthogonal decomposition theorems for mimetic finite difference methods*, SIAM J. Numer. Anal. **36** (1999), 788–909.
- [25] J. Jin, *The finite element method in electromagnetics*, John Wiley and Sons, 1993.
- [26] Jan Modersitzki, *Numerical methods for image registration*, Oxford University Press, 2004.
- [27] J. Nocedal and S.J. Wright, *Numerical optimization*, Springer, New York, 1999.
- [28] Chris. C. Paige and Michael A. Saunders, *Solution of sparse indefinite systems of linear equations*, SIAM J. Numerical Analysis **12** (1975), no. 4, 617–629.
- [29] A. Roche, *Recalage d’images médicales par inférence statistique*, Ph.D. thesis, Université de Nice, Sophia-Antipolis, France, 2001.
- [30] Torsten Rohlfing, Calvin R. Maurer, Jr., David A. Bluemke, and Michael A. Jacobs, *Volume-preserving nonrigid registration of MR breast images using free-form deformation with an incompressibility constraint*, IEEE Transactions on Medical Imaging **22** (2003), no. 6, 730–741.
- [31] D. Silvester, H. Elman, D. Kay, and A. Wathen, *Efficient preconditioning of the linearized Navier-Stokes equations*, J. Comp. Appl. Math. **128** (2001), 261–279.
- [32] U. Trottenberg, C. Oosterlee, and A. Schüller, *Multigrid*, Academic Press, 2001.
- [33] S. Turek, *Efficient solvers for incompressible flow problems*, New York, Macmillan, 1999.
- [34] Paul A. Viola, *Alignment by maximization of mutual information*, Ph.D. thesis, Massachusetts Institute of Technology, june 1995, pp. 1–155.

- [35] K.S. Yee, *Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media*, IEEE Trans. on antennas and propagation **14** (1966), 302–307.
- [36] Lei Zhu, Steven Haker, and Allen Tannenbaum, *Area preserving mappings for the visulization of medical structures*, MICCAI 2003 (2003), 277–284.