Large scale problems arising from image registration

Bernd Fischer[*1] and **Jan Modersitzki**[**]

[1] Institute of Mathematics, Wallstraße 40, University of Lübeck, Germany

Received 25 November 2004, revised 30 November 2004, accepted 2 December 2004
Published online 3 December 2004

Key words Large Scale Problems, Image Registration, Warping, Fusion, Multigrid, Optimization
MSC (2000) 35G99,49-01,65-01,65F22,65N22,65N55

Image registration, in particular medical image registration, has been subject to extensive studies in the past years. Its versatile and important applications have attracted researchers from various branches, including numerical analysts. In this note, we not only give an overview on some new and exiting approaches but also point out some of the mathematical challenges associated with these approaches. Moreover, we explain why variational based methods to image registration naturally lead to large scaled and challenging optimization and hence numerical linear algebra problems. We present various instructive examples and comment on numerical schemes.

1 Introduction

Image registration is one of today's challenging image processing problems. The objective is to find a geometrical transformation that aligns points in one view of an object with corresponding points in another view of the same object or a similar one. Particularly in medical imaging, there are many instances that demand for registration. Typical examples include the treatment verification of pre- and post-intervention images, study of temporal series of images, and the monitoring of time evolution of an agent injection subject to a patient-motion. Another important area is the need for combining information from multiple images acquired using different modalities, sometimes also called fusion. Typical examples include the fusion of computer tomography (CT) and magnetic resonance (MRI) images, of CT and positron emission tomography (PET), or of CT and ultrasound images (US). Image registration is inevitable whenever images acquired from different subjects, at different times, or from different scanners, need to be combined or compared for analysis or visualization. In the past two decades computerized image registration has played an increasingly important role in medical imaging (see, e.g., [4, 18, 34, 32, 44, 16, 27, 53, 35] and references therein).

A variety of different methods in order to solve the registration problem have been proposed. Among these are variational approaches [5, 3, 10, 28, 9, 23], partial differential equation (PDE) based methods [5, 3, 11, 14], discrete graph methods such as min-flow [41] and

[*] Corresponding author: e-mail: fischer@math.uni-luebeck.de, Phone: +0049 451 7030 434, Fax: +0049 451 7030 436
[**] e-mail: modersitzki@math.uni-luebeck.de

max-cut [1], stochastic approaches [5, 20], and dynamical programming [33, 49], to name just a few.

Due to the manifoldness of approaches, it is almost impossible to give an exhaustive overview on existing methods. Here, we decided to discuss the variational approach in more detail. There are two reasons for this choice. First of all, we believe that the variational setting is the most flexible framework and secondly, it leads to very many interesting numerical problems, in particular in numerical linear algebra.

The paper is organized as follows. We start out with a section on examples which illustrate the need for a flexible framework. This section is followed by a theoretical part, which shows on how one may phrase the different application driven demands in terms of a variational problem. In order to illustrate some of the numerical issues, we discuss in Section three in greater detail a particular application: the monitoring of breast tumor growth. Finally, we present some numerical examples which show that the presented schemes may be used to advantage in a clinical environment.

2 Examples

The following examples illustrate a variety of different demands for images registration techniques. The images have been provided by Philips Medical Systems (knee), Institute of Nuclear Medicine, RWTH Aachen (lung), University of Chicago (hands), and Lucas Center, Stanford University (breast).

Example 2.1 (Knee) This three-dimensional test case indicates the need for schemes which are able to deal with highly non-linear corrections fields. Also, this problem has about 3.2 million unknowns, which surely creates the need for state of the art numerical schemes.

Figure 1 shows CT scans of a human knee in a straight and a bent position. Part (a,b) visualize the three-dimensional data set ($128 \times 128 \times 64$ voxel) and (d,e) corresponding two-dimensional cross-sections (slice 40). The subplots (g) and (h) display the maximum intensity projection (MIP) before (g) and after (h) registration. Here, the image difference has been reduced to approximately 10%. This is also visible from the final registration result (c) and (f), respectively.

Example 2.2 (Lung) The next example is concerned with multimodal images. Here, the images display the same patient, but were acquired by different scanners. This example clearly shows, that measuring the distance between multimodal images can be a tricky and delicate matter.

Figure 2 shows a CT scan (a) and a PET image (b) of a human lung; Figure 2(c) shows the PET image after registration. Although a qualitative measurement is by no means trivial, the registered PET image appears to be much closer to the CT scan than the unregistered one. It should be noted, that the colormap of the original images has been modified in order to emphasize the differences.

Example 2.3 (Hand) Often it is desirable to guide the registration process by incorporating additional information like the location of particular anatomical landmarks. Also, depending on the specific application, it might be inevitable for the radiologist, to have a registration result which does maintain a one-to-one correspondence between these landmarks.

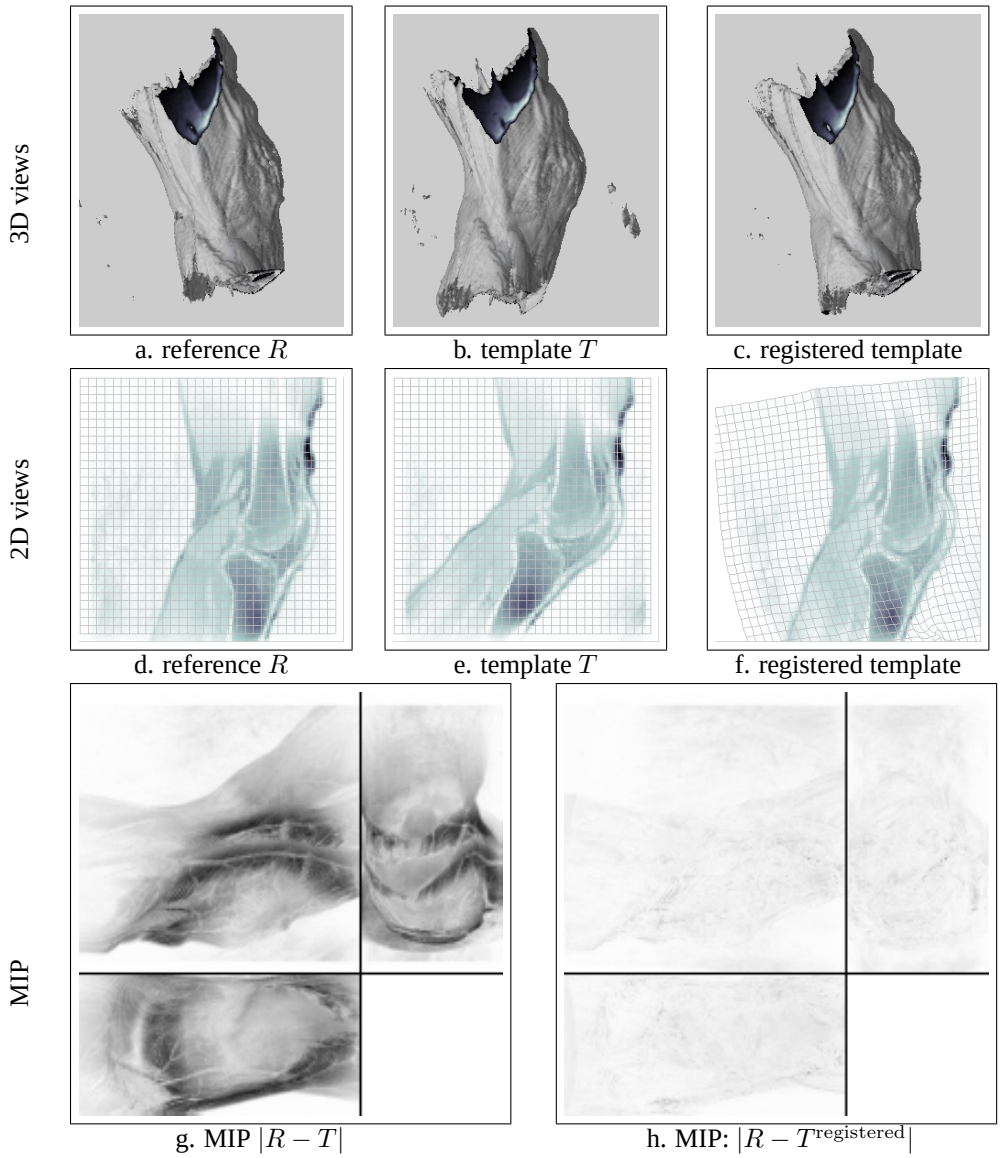
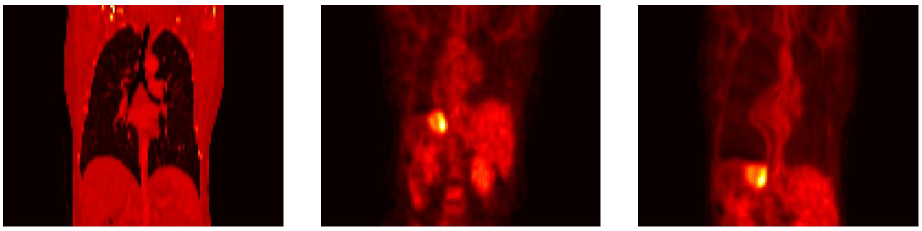


Fig. 1 Registration results for two $128 \times 128 \times 64$ MRI scans of a human knee; 3D views of R (a), T (b), and $T^{\mathrm{registered}}$ (c); the generic slice 40 with grid for R (d), T (e), and $T^{\mathrm{registered}}$ (f); maximum intensity projections for $|R-T|$ (g) and $|R-T^{\mathrm{registered}}|$ (h).

Figure 3 shows a typical test example along this lines. Here, two modified X-rays of a human hand (a,b) have been equipped with six particular anatomical landmarks. The subplot (c) displays the final registration, where a one to one correspondence of the landmarks is guaranteed.



a. reference R b. template T c. registered template

Fig. 2 CT image of a human lung (a), PET images before (b) and after (c) registration (images modified for visualization purposes).

a. reference R b. template T c. registered template

Fig. 3 Modified X-rays of a human hand with anatomical landmarks

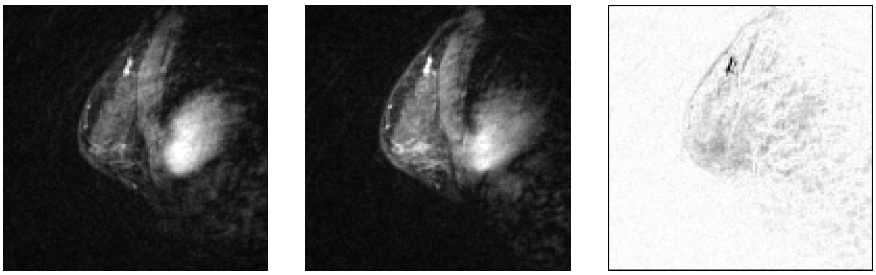
Example 2.4 (Breast) An application of particular clinical interest is the registration of pairs of images acquired before and after contrast administration. Here, the trouble is, that some regions tend to enlarge due to the contrast uptake but not due to a possible patient movement. An unconstrained registration technique then would be trapped by this enlargement, which may result in misleading registration results.

The following figure shows a challenging example. Here, the overall goal is tumor monitoring. Therefore a change of tissue volume can become critical even if it results in much more pleasing visual results.

Figure 4 shows two two-dimensional slices of an MR scan of a female breast taken during contrast agent uptake (a,b). The difference image (c) shows some suspicious regions, which are easily identified since the images are registered. This example is discussed in detail in Section 4.2, where also numerical results are shown.

3 The registration problem

The previous examples indicate the need for a broad and general mathematical model, capable of dealing with a variety of situations, easy to modify and to extend. There are various approaches to achieve this goal. However, a discussion of all possible techniques is far beyond the scope of this paper. Here, we focus on a model based on a variational approach. This is for basically two reasons. First of all, the continuous setting enables one to use different image



a. reference R b. template T c. difference

Fig. 4 MRI's of a female breast, LEFT: during the wash-in phase, MIDDLE: during the wash-out phase, and RIGHT: difference image.

resolutions or equivalently different discretizations, which are naturally linked by the underlying continuous problem. Secondly, as we will see in a moment, the image registration problem can be treated as an optimization problem. For the applications of interest, the problem is typically large scaled and in many applications also time critical. Therefore, a successful scheme depends on efficient optimization techniques which generally rely on derivatives, which may be hard to calculate in a discrete setting.

To begin with, we present some notations. Given a *reference* image R and a *template* image T, the goal of image registration is to find a "*reasonable*" transformation such that the "*distance*" between the reference image and the deformed template image is small. With $d \in \mathbb{N}$ we denote the spatial dimension of the given images $R, T : \mathbb{R}^d \to \mathbb{R}$, which are assumed to be sufficiently smooth (see the discussion in Section 4.3.4). Thus, $T(\boldsymbol{x})$ is the gray value at a spatial position $\boldsymbol{x}$. Moreover, without loss of generality, we assume that the contents of the images are contained in a bounded domain $\Omega = (0,1)^d$.

The goal is to find a deformation $\boldsymbol{u} : \mathbb{R}^d \to \mathbb{R}^d$ such that the deformed image $T[\boldsymbol{u}]$, with $T[\boldsymbol{u}](\boldsymbol{x}) := T(\boldsymbol{x} + \boldsymbol{u}(\boldsymbol{x}))$, is "*similar*" to R.

Next, we state our general registration model. Note, that the building blocks are easily exchangeable and may be tuned with respect to the application in question.

Image registration: Given two images R and T, find a deformation $\boldsymbol{u}$, such that

$$\mathcal{D}[R, T; \boldsymbol{u}] + \alpha \mathcal{S}[\boldsymbol{u}] + \beta \mathcal{C}^{\text{soft}}[\boldsymbol{u}] = \min \tag{1a}$$

$$\text{subject to } \mathcal{C}^{\text{hard}}[\boldsymbol{u}](\boldsymbol{x}) = 0 \quad \text{for all} \quad x \in \Omega_{\mathcal{C}}. \tag{1b}$$

Here, $\mathcal{D}$ denotes a typically intensity based *distance measure*, $\mathcal{S}$ a *regularizer*, $\mathcal{C}^{\text{soft}}$ some soft constraints (penalization), and $\mathcal{C}^{\text{hard}}$ some hard constraints. The number $\alpha > 0$ is a regularization parameter, the number $\beta \geq 0$ a weighting factor, and the set $\Omega_{\mathcal{C}}$ is a subset of the image domain. On the first glance it might appear peculiar, that we do distinguish between hard- and soft constraints. Note, however, that the hard constraints are to be fulfilled by the minimizer, whereas the soft constraints, depending on the parameter β, constitute a relaxed version of the hard constraints. For example, when the application involves some

user-prescribed landmarks, one may want to consider only soft constraints, as the landmarks itself are located approximately and exhibit some uncertainty (see, also Example 3.3). On the other hand, in tumor growth monitoring a hard volume preserving constraint is inevitable.

Existence and uniqueness of different realizations of the problem are discussed for example in [8] and [35]. In this paper we focus on numerical issues. Additional building blocks are discussed for example in [11, 38, 29, 31, 39, 15]. A typical explicit regularization is given by restricting the wanted displacement to a certain space, for example such that the transformation is affine linear or belongs to a certain spline space. For this so-called parametric registration, see, e.g., [32, 35].

Choices for the building blocks $\mathcal{D}$, $\mathcal{S}$, and $\mathcal{C}^{\mathrm{hard,soft}}$ as well as the choices for parameters are typically application dependent and normally involve an extended discussion with the clinical partner. The following examples illustrate some specific applications and concretizations which have been proven satisfactorily. An objective validation technique for registration results is yet to be found.

Example 3.1 (Knee – continued) We continue the knee Example 2.1. Here, the images are three-dimensional, $d = 3$. The CT scans of the human knee have the same image modality. Therefore our distance measure $\mathcal{D}$ is based on a direct comparison of image intensities (*sum of squared differences, SSD*),

$$\mathcal{D}^{\mathrm{SSD}}[R, T; \boldsymbol{u}] = \tfrac{1}{2} \int_\Omega \big(T(\boldsymbol{x} + \boldsymbol{u}(\boldsymbol{x})) - R(\boldsymbol{x})\big)^2 \, d\boldsymbol{x}. \tag{2}$$

As a regularizer, we used the elastic potential

$$\mathcal{S}^{\mathrm{elas}}[\boldsymbol{u}] = \tfrac{1}{2} \int_\Omega \left\| \mathcal{B}^{\mathrm{elas}}[\boldsymbol{u}] \right\|^2_{\mathbb{R}^d} \, d\boldsymbol{x} \tag{3}$$

where

$$\mathcal{B}^{\mathrm{elas}}[\boldsymbol{u}] = \sqrt{\mu}\Big((\nabla u_1)^\top, \ldots, (\nabla u_d)^\top, \sqrt{2 + \lambda/\mu}\, \nabla \cdot \, \boldsymbol{u} \Big)^\top,$$

μ and λ are the so-called Lamé-constants. This choice was motivated by the goal to ensure a smooth (elastic) transition from the template to the reference image. Here, we do not introduce additional constraints: $\beta = 0$, $\Omega_{\mathcal{C}} = \emptyset$.

Example 3.2 (Lung – continued) Obviously a direct comparison of image intensities is not advisable for Example 2.2. Here, we use *mutual information* as a distance measure. Roughly speaking, mutual information measures the amount of information of R contained in T; see, e.g., [7, 50, 35] and references therein. Let the joint density of the gray value distribution of R and T be denoted by ρ,

$$\rho^{(R,T)}(r, t) = \int_\Omega \chi\{\boldsymbol{x} \mid R(\boldsymbol{x}) = r \wedge T(\boldsymbol{x}) = t\} \, d\boldsymbol{x}, \tag{4}$$

where χ is the characteristic function, and the individual gray value distribution by

$$\rho^R(r) = \int_{\mathbb{R}} \rho^{(R,T)}(r, t) \, dt, \quad \rho^T(t) = \int_{\mathbb{R}} \rho^{(R,T)}(r, t) \, dr.$$

Then, mutual information is defined by

$$\mathcal{D}^{\mathrm{MI}}[R, T; \boldsymbol{u}] = \int_{\mathbb{R}^2} \rho^{(R,T[\boldsymbol{u}])} \log_2 \frac{\rho^{(R,T[\boldsymbol{u}])}}{\rho^R \cdot \rho^{T[\boldsymbol{u}]}} \, d(r, t).$$

As a regularizer $\mathcal{S}^{\mathrm{curv}}$ we use the curvature regularizer [14],

$$\mathcal{S}^{\mathrm{curv}}[\boldsymbol{u}] = \tfrac{1}{2}\int_\Omega \|\mathcal{B}^{\mathrm{curv}}[\boldsymbol{u}]\|^2_{\mathbb{R}^d}\; d\boldsymbol{x}$$

with

$$\mathcal{B}^{\mathrm{curv}}[\boldsymbol{u}] = \Delta\boldsymbol{u} = (\Delta u_1, \ldots, \Delta u_d)^\top. \tag{5}$$

This regularizer has been proven to produce very satisfying results, even when a large transformation is necessary. No constraints are added.

Example 3.3 (Hand – continued) For Example 2.3, we use the SSD based distance measure (2) and the curvature regularizer (5). In contrast to the previous examples, we add additional hard constraints. Let $\boldsymbol{t}^j$ and $\boldsymbol{r}^j$, $j = 1, \ldots, m$, be some user prescribed landmarks. The goal is to specify $\boldsymbol{u}$ such that

$$\boldsymbol{t}^j + \boldsymbol{u}(\boldsymbol{t}^j) = \boldsymbol{r}^j, \quad j = 1, \ldots, m.$$

Using the d-dimensional point evaluation functional $\delta_{\boldsymbol{x}}$ these constraints can also be phrased as $\mathcal{C}_j[\boldsymbol{u}] = \boldsymbol{r}^j + \delta_{\boldsymbol{t}^j}[\boldsymbol{u}] - \boldsymbol{t}^j$, and hence the landmark constraints are given by

$$\mathcal{C}^{\mathrm{landmarks}} = (\mathcal{C}_1, \ldots, \mathcal{C}_m)^\top, \quad \Omega_{\mathcal{C}} = \Omega.$$

Note that these hard constraints imply that our transformation produces an exact one-to-one correspondence of the landmarks. However, in many application is it impossible to determine landmarks exactly; see, e.g. [40]. Therefore it is an interesting option to replace the hard landmarks by soft landmarks, which may be phrased as

$$\mathcal{C}^{\mathrm{soft}}[\boldsymbol{u}] = \tfrac{1}{2}\sum_{j=1}^{m}\Big(\boldsymbol{t}^j + \boldsymbol{u}(\boldsymbol{t}^j) - \boldsymbol{r}^j\Big)^2.$$

Example 3.4 (Breast – continued) Following [24], we use SSD (2) as a distance measure and the elastic potential (3) as a regularizer for Example 2.4. Hard constraints enforce volume preservation and are based on the determinant of the Jacobian of the transformation $\boldsymbol{x} + \boldsymbol{u}(\boldsymbol{x})$,

$$\mathcal{C}^{\mathrm{VP}}[\boldsymbol{u}](\boldsymbol{x}) = \det(I_d + \nabla\boldsymbol{u}(\boldsymbol{x})) - 1 = 0 \quad \text{for all} \quad \boldsymbol{x} \in \Omega_{\mathcal{C}}, \tag{6}$$

where for simplicity $\Omega_{\mathcal{C}} = \Omega$. Note that the set $\Omega_{\mathcal{C}}$ could easily be based on a segmentation of the image. For example, we could ask for volume preservation only for tissue and/or bones.

In this example, soft constraints

$$\mathcal{C}^{\mathrm{soft}}[\boldsymbol{u}] = \int_{\Omega_{\mathcal{C}}} (\mathcal{C}^{\mathrm{VP}}[\boldsymbol{u}])^2\; d\boldsymbol{x}$$

are not advisable. Firstly, after an overall minimization of the registration problem, we cannot expect $\mathcal{C}^{\mathrm{soft}}[\boldsymbol{u}^*] = 0$ for the solution $\boldsymbol{u}^*$. Secondly, and most importantly, even if $\mathcal{C}^{\mathrm{soft}}[\boldsymbol{u}^*]$ is small, the integrant $\mathcal{C}^{\mathrm{VP}}[\boldsymbol{u}^*]$ can be quite large in some small regions like, for example, a tumor.

4 Numerical schemes

Having introduced some typical applications, we will now turn our attention to the actual implementation of a registration problem. To begin with, we discuss the breast example in more detail. This application has also been studied in [39, 24]. Here, we highlight some mathematical and implementation issues, which came up in the course of solving this complex registration problem. Moreover, we present some concrete numerical results and some performance measurements and conclude this section with remarks concerning the general registration problem (1).

4.1 Numerics for the breast example

With $\mathcal{D} = \mathcal{D}^{\mathrm{SSD}}$ (cf. (2)), $\mathcal{S} = \mathcal{S}^{\mathrm{elas}}$ (cf. (3)), $\beta = 0$, $\mathcal{C} = \mathcal{C}^{\mathrm{VP}}$ (cf. (6)), and $\Omega_{\mathcal{C}} = \Omega$, the continuous registration problem reads

$$\mathcal{D}[R, T; \boldsymbol{u}] + \alpha \mathcal{S}[\boldsymbol{u}] = \min \tag{7a}$$

$$\text{subject to } \mathcal{C}[\boldsymbol{u}](\boldsymbol{x}) = 0 \quad \text{for all} \quad x \in \Omega. \tag{7b}$$

Depending on the actual data, the choice of the regularization parameter α might be critical for the success of the scheme. This issue will be briefly discussed in Section 4.5.

In the next subsections, we will highlight some of the details concerning a proper implementation of the above problem. For the ease of presentation, we restrict the discussion to the two-dimensional case and refer to [24] for higher dimensions and for further details.

4.1.1 Discretization

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 so-called LBB conditions [2, 17]. Further complications arise in this case where we need to discretize differential operators such as the divergence and the curl. A conservative compact discretization schemes could be achieved by either mixed finite elements or by staggered grids. Here we focus on staggered grids; see [24] for an extended discussion.

The discrete version of $\boldsymbol{u}$ is denoted by $\vec{\boldsymbol{u}} = ((\vec{\boldsymbol{u}}^1)^\top, (\vec{\boldsymbol{u}}^2)^\top)^\top$, where $\vec{\boldsymbol{u}}^1$ is collected on the ▼-grid and $\vec{\boldsymbol{u}}^2$ on the ▶-grid, respectively; see Figure 5. Note that it is common in imaging to have the first coordinate top-down and the second left-right.

Partial derivatives are approximated be centered finite differences and thus normal derivative are naturally located at the cell center positions (•). For tangential direction we imposed Neumann boundary conditions. Using the discrete analogs ∂_j^h, ∇_j^h, and $\nabla^h \cdot$ for the partial differential operators ∂_j, gradient ∇, and divergence $\nabla \cdot$, the elastic potential $\mathcal{S}^{\mathrm{elas}}$ (3) is discretized as

$$\mathcal{S}^h(\vec{\boldsymbol{u}}) = \|\boldsymbol{B}\vec{\boldsymbol{u}}\|_2^2 := \frac{\lambda+\mu}{2} \| \nabla^h \cdot \vec{\boldsymbol{u}}\|^2 + \frac{\mu}{2}(\| \nabla_1^h \vec{\boldsymbol{u}}^1\|^2 + \| \nabla_2^h \vec{\boldsymbol{u}}^2\|^2).$$

Upon differentiation with respect to $\vec{\boldsymbol{u}}$, we obtain the Navier-Lamé operator

$$\mathcal{S}_{\vec{\boldsymbol{u}}}^h(\vec{\boldsymbol{u}}) = (\lambda+\mu)(\nabla^h \cdot)^\top \nabla^h \cdot \vec{\boldsymbol{u}} - \mu \Delta_h \vec{\boldsymbol{u}} =: \boldsymbol{A}\vec{\boldsymbol{u}},$$

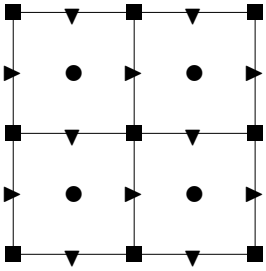


Fig. 5 Staggered grids (for four pixels), nodal ■, cell centered •, face staggered grids (▼ in x_1-, ► in x_2-direction).

where Δ_h denotes the usual five point discrete vector Laplacian.

Since we aim for a differentiable objective function, we embed the image data into a smooth spline space. With the projection operator P_j from the jth staggered grid (▼, ►) to the cell centered grid (•), we set

$$T(\vec{\boldsymbol{u}}) := T^{\mathrm{spline}}(\vec{\boldsymbol{x}}^1 + P_1\vec{\boldsymbol{u}}^1, \vec{\boldsymbol{x}}^2 + P_2\vec{\boldsymbol{u}}^2),$$

where T^{spline} denotes the appropriate chosen spline approximation. Now, we are able to safely compute the Jacobian of T via

$$T_{\vec{\boldsymbol{u}}}(\vec{\boldsymbol{u}}) := \tfrac{\partial T}{\partial \vec{\boldsymbol{u}}}(\vec{\boldsymbol{u}}) = \Big(\mathrm{diag}(P_1^\top \partial_1 T), \mathrm{diag}(P_2^\top \partial_2 T)\Big), \tag{8}$$

where the partial derivatives $\partial_j T$ are evaluated at the spatial positions $(\vec{\boldsymbol{x}}^1 + P_1\vec{\boldsymbol{u}}^1, \vec{\boldsymbol{x}}^2 + P_2\vec{\boldsymbol{u}}^2)$.

A discretization of $\mathcal{D}^{\mathrm{SSD}}$ is straightforward,

$$\boldsymbol{D}(\vec{\boldsymbol{u}}) := \tfrac{1}{2}\|T(\vec{\boldsymbol{u}}) - R\|_2^2 \quad \text{and thus} \quad \boldsymbol{D}_{\vec{\boldsymbol{u}}}(\vec{\boldsymbol{u}}) = T_{\vec{\boldsymbol{u}}}(\vec{\boldsymbol{u}})^\top (T(\vec{\boldsymbol{u}}) - R).$$

Discretizing the constraint $\mathcal{C}$ involves a polynomial in the derivatives of $\vec{\boldsymbol{u}}$. With $\odot$ denoting the Hadamard (pointwise) multiplication, we use

$$\boldsymbol{C}(\vec{\boldsymbol{u}}) = \partial_1^h \vec{\boldsymbol{u}}^1 + \partial_2^h \vec{\boldsymbol{u}}^2 \; + \; \partial_1^h \vec{\boldsymbol{u}}^1 \odot \partial_2^h \vec{\boldsymbol{u}}^2 - (P\partial_2^h \vec{\boldsymbol{u}}^1) \odot (P\partial_1^h \vec{\boldsymbol{u}}^2),$$

where $P = P_1 P_2$ is a projector from the nodal to the cell centered grid. The derivatives $\boldsymbol{C}_{\vec{\boldsymbol{u}}}$ are then calculated as follows,

$$\boldsymbol{C}_{\vec{\boldsymbol{u}}^1}(\vec{\boldsymbol{u}}) = \big(I + \mathrm{diag}(\partial_2^h \vec{\boldsymbol{u}}^2)\big)\partial_1^h + \mathrm{diag}(P\partial_1^h \vec{\boldsymbol{u}}^2) P \partial_2^h, \tag{9a}$$

$$\boldsymbol{C}_{\vec{\boldsymbol{u}}^2}(\vec{\boldsymbol{u}}) = \big(I + \mathrm{diag}(\partial_1^h \vec{\boldsymbol{u}}^1)\big)\partial_2^h + \mathrm{diag}(P\partial_2^h \vec{\boldsymbol{u}}^1) P \partial_1^h. \tag{9b}$$

4.1.2 Optimization

We are now ready to phrase the discrete analog of the image registration problem (7),

$$\boldsymbol{J}(\vec{\boldsymbol{u}}) := \boldsymbol{D}(\vec{\boldsymbol{u}}) + \alpha \boldsymbol{S}(\vec{\boldsymbol{u}}) = \min \tag{10a}$$

$$\text{subject to} \quad \boldsymbol{C}(\vec{\boldsymbol{u}}) = 0. \tag{10b}$$

In order to solve this problem numerically, we use the SQP framework; see [36, 24] for a detailed discussion. To this end, let $\vec{\boldsymbol{p}}$ be a cell-centered vector of Lagrange multipliers. The Lagrangian of the problem then looks like

$$\boldsymbol{L}(\vec{\boldsymbol{u}},\vec{\boldsymbol{p}}) \quad = \quad \boldsymbol{D}(\vec{\boldsymbol{u}}) + \tfrac{\alpha}{2}\left\|\boldsymbol{B}\vec{\boldsymbol{u}}\right\|^2 + \boldsymbol{C}(\vec{\boldsymbol{u}})^\top \vec{\boldsymbol{p}}.$$

Differentiating with respect to $\vec{\boldsymbol{u}}$ and $\vec{\boldsymbol{p}}$, we obtain the following discrete Euler-Lagrange equations

$$0 \quad = \quad \boldsymbol{L}_{\vec{\boldsymbol{u}}}(\vec{\boldsymbol{u}},\vec{\boldsymbol{p}}) = \boldsymbol{D}_{\vec{\boldsymbol{u}}}(\vec{\boldsymbol{u}}) + \alpha\boldsymbol{B}^\top\boldsymbol{B}\vec{\boldsymbol{u}} + \boldsymbol{C}_{\vec{\boldsymbol{u}}}(\vec{\boldsymbol{u}})^\top\vec{\boldsymbol{p}}, \tag{11a}$$

$$0 \quad = \quad \boldsymbol{L}_{\vec{\boldsymbol{p}}}(\vec{\boldsymbol{u}},\vec{\boldsymbol{p}}) = \boldsymbol{C}(\vec{\boldsymbol{u}}). \tag{11b}$$

To solve this non-linear system numerically, we apply a Newton-type method. The idea is to approximate the (1,1) block of the Hessian by

$$\boldsymbol{H} := \alpha\boldsymbol{B}^\top\boldsymbol{B} + T_{\vec{\boldsymbol{u}}}^\top T_{\vec{\boldsymbol{u}}}, \tag{12}$$

where $T_{\vec{\boldsymbol{u}}}$ is defined in (8). Thus, the following linear system of equations

$$\begin{pmatrix} \boldsymbol{H} & \boldsymbol{C}_{\vec{\boldsymbol{u}}}^\top \\ \boldsymbol{C}_{\vec{\boldsymbol{u}}} & 0 \end{pmatrix} \begin{pmatrix} s_{\vec{\boldsymbol{u}}} \\ s_{\vec{\boldsymbol{p}}} \end{pmatrix} = - \begin{pmatrix} \boldsymbol{L}_{\vec{\boldsymbol{u}}}(\vec{\boldsymbol{u}},\vec{\boldsymbol{p}}) \\ \boldsymbol{L}_{\vec{\boldsymbol{p}}}(\vec{\boldsymbol{u}},\vec{\boldsymbol{p}}) \end{pmatrix} =: \begin{pmatrix} \boldsymbol{r} \\ \boldsymbol{s} \end{pmatrix} \tag{13}$$

has to be solved at each iteration. The system (13) 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., [21, 48]) and the solution for this case has been addressed by many authors; see, e.g., [48, 43, 47]. Here, we have used MINRES [37] with a block diagonal preconditioner, as proposed in [43]. This preconditioner may be written as

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

where $\hat{\boldsymbol{S}}$ is an approximation to the Schur complement $\boldsymbol{S} := \boldsymbol{C}_{\vec{\boldsymbol{u}}}\,\boldsymbol{H}^{-1}\,\boldsymbol{C}_{\vec{\boldsymbol{u}}}^\top$. There exist many options for a sound choice of $\hat{\boldsymbol{S}}$. Here, we use the approximation suggested in [43],

$$\hat{\boldsymbol{S}}^{-1} = (\boldsymbol{C}_{\vec{\boldsymbol{u}}}\,\boldsymbol{C}_{\vec{\boldsymbol{u}}}^\top)^{-1}\,\boldsymbol{C}_{\vec{\boldsymbol{u}}}\,\boldsymbol{H}\,\boldsymbol{C}_{\vec{\boldsymbol{u}}}^\top\,(\boldsymbol{C}_{\vec{\boldsymbol{u}}}\,\boldsymbol{C}_{\vec{\boldsymbol{u}}}^\top)^{-1}. \tag{14}$$

Note that $(\boldsymbol{C}_{\vec{\boldsymbol{u}}}\,\boldsymbol{C}_{\vec{\boldsymbol{u}}}^\top)^{-1}\,\boldsymbol{C}_{\vec{\boldsymbol{u}}}$ is the pseudo-inverse of $\boldsymbol{C}_{\vec{\boldsymbol{u}}}$. The application of the preconditioner only involves a multiplication of $\boldsymbol{H}^{-1}$ and $\hat{\boldsymbol{S}}^{-1}$ with a vector, which may be accomplished by multigrid techniques as discussed in [23].

4.2 Results for the breast example

The above scheme has been implemented using MATLAB 6.5. The following computations are performed on a DELL Inspiron 8600 notebook (1.4 GHz, 1GB RAM, 60G disk space) under Windows XP. The computation time for the breast Example 2.4 and 3.4 (128×128 images) is about 296 seconds for the unconstrained registration (68 iteration needed to fulfill our stopping criterion $\left\|\vec{\boldsymbol{u}}^{(k-1)} - \vec{\boldsymbol{u}}^{(k)}\right\| \leq \mathrm{tol}_U := 10^{-1}$, for the constrained problem we spend about 846 seconds (23 iteration needed).

The scheme has been applied to the images shown in Fig. 4. In Fig. 6 the results after two (2nd row) and ten iterations (3rd row) of the unconstrained registration as well as after ten iterations of the volume preserving (VP) constrained registration (4th row) are shown. The numerical results are summarized in Table 1. After ten iteration both schemes have converged.

Table 1 Numerical results for the un- and VP-contrained registrations; k is the number of performed iterations.

		α	k	$\boldsymbol{D}(\vec{\boldsymbol{u}}^{(k)})/\boldsymbol{D}(0)$	$\|\boldsymbol{C}(\vec{\boldsymbol{u}}^{(k)})\|_\infty$
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 satisfactory. The explanation is that small spots which are related to noise in the MRI images and are 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.

In Fig. 7, we display the pointwise map of the change of volume. Using the unconstrained approach, we observe a considerable change of volume for the breast with a peak value of 1.36. Thus, part of the breast has been enlarged by a factor of 2.36. For the constrained approach, we observe that the volume change is below a user supplied threshold (here, $\mathrm{tol}_C = 10^{-6}$) everywhere. In fact, since we used a quasi-Newton scheme for the projection, the numbers are around 10^{-9}.

4.3 The general case

Although the numerical scheme, outlined in the previous section, is tuned for a specific application, we can easily identify some general patterns. In order to simplify the discussion, we concentrate on the unconstrained registration problem.

4.3.1 Problem size

When working with real life registration problems, on has to be aware of the fact, that the problem size can become quite large. For example: for the very modest example of a volume preserving two-dimensional registration of two 128×128 images, the matrix $\boldsymbol{H}$ (cf. (12)) is $N \times N$ where $N \approx 32 \cdot 10^3$ and $\boldsymbol{C}_{\vec{\boldsymbol{u}}}$ (cf. (9)) is about$N/2 \times N$; for the still modest three-dimensional knee example we have $N \approx 3 \cdot 10^6$. However, there are many instances in a clinical environment, were a real-time registration is needed. Among those is the so-called brain-shift problem. Here, the registration should compensate for the shift of brain in course of a brain surgery; see, e.g. [45, 46] and references therein. For this demanding example $N \approx 50 \cdot 10^6$, whereby the allowed practical solution time is restricted to a few minutes. Therefore, more complex schemes, which may please the soul of a mathematician,

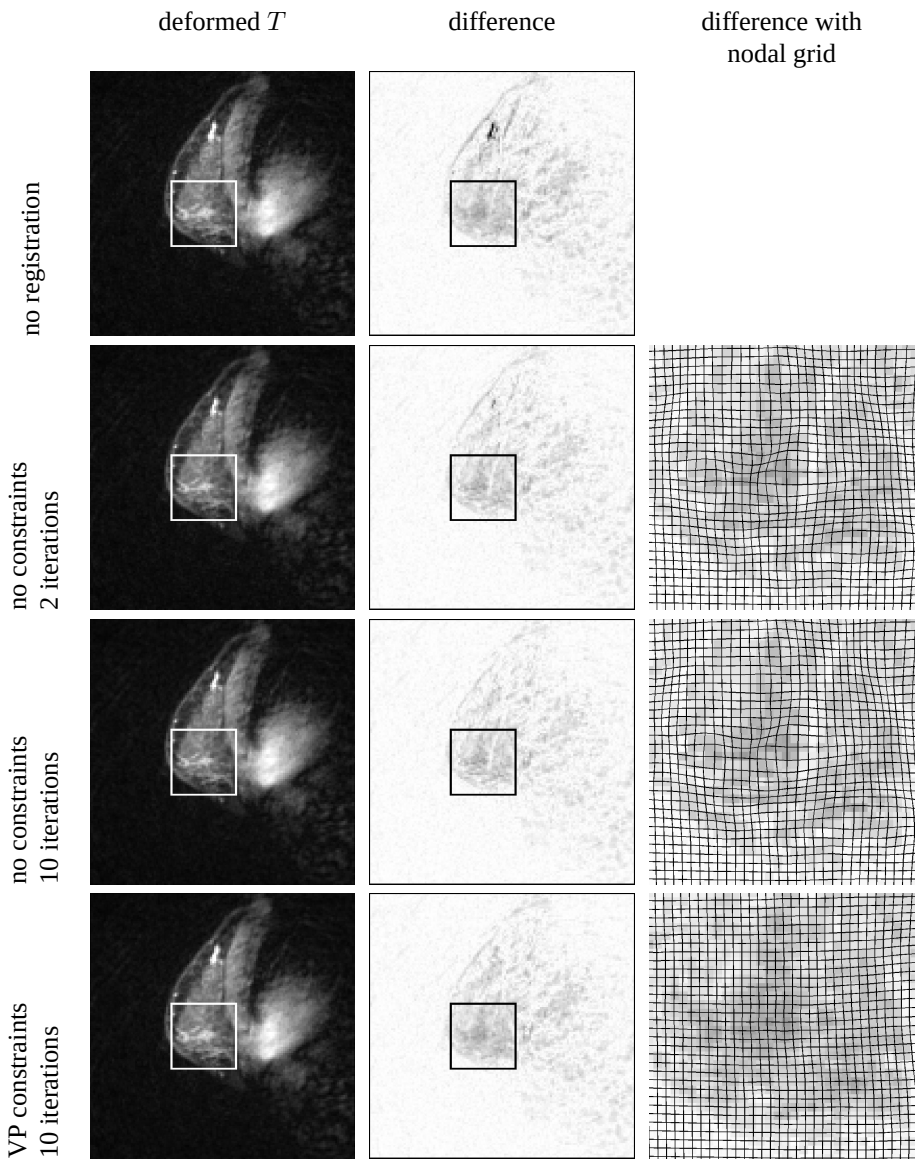


Fig. 6 Registration results for the images of Fig. 4. LEFT COLUMN deformed template images $T(\vec{\boldsymbol{u}})$, MIDDLE COLUMN difference image $|R - T(\vec{\boldsymbol{u}})|$ with region of interest (ROI), RIGHT COLUMN ROI with nodal grid, vertices connected by straight lines ; ROW 1: no registration,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

are practically relevant only if they allow for a fast implementation. Also, there has been considerable research towards parallel solution techniques for real-life applications; see, e.g. [52].

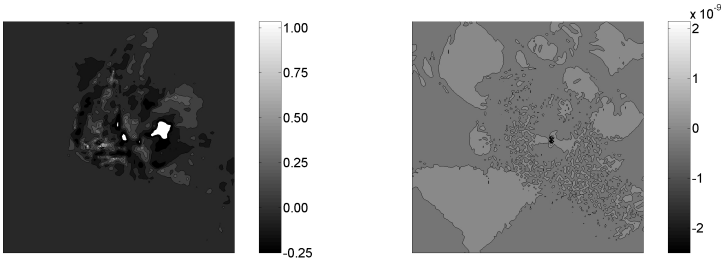


Fig. 7 Volume preservation of the unconstrained (LEFT) and constrained (RIGHT) registration results for the MRI example.

4.3.2 Parametric registration

Another discretization option is to use a finite expansion of $\boldsymbol{u}$ in terms of some basic functions. A very prominent example in this direction are splines; see e.g. [42]. This approach may be seen as a specific regularization technique, as the solution inherits the smoothness of the chosen basis functions. Another often used example for such a so-called parametric registration [32, 35] is the so-called affine linear registration. Here, $\boldsymbol{u} = Q\boldsymbol{z}$, where Q is a collection of basis functions and $\boldsymbol{z}$ denotes the parameter set. For this situation, one may set $\alpha = 0$ and in absence of constraints, Eq. (13) simplifies considerably:

$$\boldsymbol{Q}^\top \boldsymbol{H}\boldsymbol{Q}\delta_{\boldsymbol{z}} = \boldsymbol{Q}^\top \boldsymbol{r}.$$

Due to its effectiveness and simplicity, this approach is often used in practice. It may been even seen as todays industry standard. However, its applicability is restricted by its linear nature and does often fail to produce satisfactory results. Nevertheless, it is frequently used as a preprocessing step for nonlinear technique.

4.3.3 Full multigrid versus Newton type techniques

As for many problems in numerical analysis, there exist various ways to attack the non-linear problem (1). For example, a non-linear full multigrid technique might be applied; see, e.g., [28, 9]. However, we found a splitting into the Newton-type outer iteration and an inner iteration for the linear systems to be more flexible. One reason is the fact that this choice offers the possibility of applying different, application dependent, solution techniques for the linear system like, e.g., multigrid [26, 47], conjugate gradient [30], MINRES or SYMMLQ [37], or even direct schemes.

Another reason for the splitting approach is its capability of dealing with different application demands, that is, its main building blocks are easily exchangeable. For example, changing the regularizer, simply results in a different matrix $\boldsymbol{B}$. Replacing the distance measure $\mathcal{D}^{\mathrm{SSD}}$ by an alternative like mutual information does only change the Hessian (12). However, care has to be taken, since derivatives of the joint density $\rho^{(R,T(\vec{\boldsymbol{u}}))}$ (cf. (4)) need to be computed; see [25] for details. Also, different constraints are straightforward to incorporate; see e.g. [13] for landmark constraints and [22] for a space splitting idea, respectively.

4.3.4 Discretize-optimize versus optimize-discretize

The presented scheme is a so-called discretize-optimize approach. That is, the continuous problem is first discretized and then the obtained optimization problem is solved. It is this underlying optimization problem, which makes this approach very attractive. Not only efficient optimization techniques may be used to advantage, but also its variational form setting, the Euler-Lagrange equations (13), open the door to efficient solution techniques for partial differential equation; see [24] for an extended discussion.

To arrive at a sound discretization of the objective function in the approach above, it has to be expressed in terms of continuously differentiable functions, say splines. The price paid for this requirement is high, as the computation of spline coefficients and in particular the evaluation of spline functions is time consuming. A computational attractive alternative can be based on the discretization of the continuous Euler-Lagrange equation of the unconstrained registration problem, which is known as optimize-discretize approach. A minimizer is characterized by

$$\boldsymbol{f}(R, T; \boldsymbol{u}) + \mathcal{A}[\boldsymbol{u}] = \mathbf{0}, \tag{15}$$

where the so-called force-field $\boldsymbol{f}$ if the Gâteaux-derivative of the distance measure and $\mathcal{A} = \mathcal{B}^*\mathcal{B}$ is related to the regularizer. Fast schemes are based on a fixed-point type iteration $\mathcal{A}[\boldsymbol{u}^{(k+1)}] = f(\boldsymbol{u}^{(k)})$, $\boldsymbol{u}^{(0)} \equiv 0$, where all computational extensive parts are hidden in the forward problem, the computation of $\boldsymbol{f}(\boldsymbol{u}^{(k)})$. These schemes are of particularly interest, if a discretization of the partial differential operator is highly structured. A typical example is the diffusion regularizer, which allows for a direct solution technique in linear time $\mathcal{O}(n)$; see [11, 35] for details. The computation of image derivatives is based on finite differences and the computation of the deformed image is based n a d-linear interpolation. Hence, a computation of spline coefficients and an evaluation of spline functions is avoided. This can speed up the computations tremendously.

4.3.5 Scale space

Sometimes it is desirable to register the given images subsequently within different scales. Here, the spline based approach provides a natural scale parameter. The idea is to approximate the images with a varying degree of exactness. A poor approximation should only capture the main image features and may be viewed as a smooth version of the image. Due to the simplicity of the smoothed images, the corresponding registration is easy to accomplish. This result then may be used as starting guess for approximations which capture more and more details. The smoothing parameter (degree of exactness) might be used as a degradation parameter or be chosen by some optimality criterion like, e.g. the Generalized Cross Validation (GCV) method [19]. For data interpolation using B-splines see [51].

4.4 Multi-resolution

In addition to the techniques discussed in Section 4.3.5, multi-resolution type approaches like for example image pyramids are often used. One obvious reason is that multi-resolution techniques typically reduce the computation time. Major corrections are computed on a computationally inexpensive coarse grid, on the expensive finer grids only a few corrections steps

are needed. A second major advantage is that these strategies prevent to converge to local minima.

4.5 The parameter choice

In contrast to many other ill-posed problems, where efficient strategies like GCV are available to automatically estimate the regularization parameter, for image registration satisfactory strategies are yet missing. The typical choice and the validation of the registration results are typically performed by the trained experts.

5 Summary

We present a general approach to image registration based on a variational formulation, where the overall goal is to minimize a certain energy functional. However, it is well-known that image registration is an ill-posed problem. Therefore, for a particular given application, one tries to add as much pre-information as possible to the functional. One step in this direction is to add a designed regularizer, which on one hand penalizes unwanted solutions and on the other hand is inevitable to make the problem well-posed. In this paper, we consider a further step, which is to provide additional information by adding additional constraints. Constraints are sometimes treated by penalization [6, 39], which in a sense is similar to the regularized approach. However, in the penalized version it is not guaranteed that the constraints hold whereas in the constrained approach it is guaranteed.

We also present an implementation of this general approach for volume preserving constraints. In particular, we present a stable and consistent staggered grid discretization of the continuous variational approach. We show results for a real life application (monitoring of tumor growth in a female breast) and discuss performance items.

From the numerical point of view, the constrained approach leads to a very challenging and highly non-linear optimization problem where both the discretization and the numerical optimization are not obvious. We used state-of-the-art SQP optimization techniques that allow for an efficient solution of the registration problem. Our formulation opens up a few avenues of research such as effective solvers for the KKT systems and inexact SQP methods.

Finally, we comment on various features of modern registration techniques and outline some of the main research streams in image registration.

Acknowledgements Jan Modersitzki was supported by the US National Institutes of Health under Grant NIHR01 HL 068904. We are indebted to Bruce Daniels, Eldad Haber, and Stefan Heldmann for providing exciting data and valuable discussions.

References

[1] Y. BOYKOV, O. VEKSLER, AND R. ZABIH, *Fast approximate energy minimization via graph cuts*, in Proceedings of the International Conference on Computer Vision, 1999.
[2] F. BREZZI AND M. FORTIN, *Mixed and hybrid finite element methods*, Springer, 1991.
[3] M. BRO-NIELSEN, *Medical Image Registration and Surgery Simulation*, PhD thesis, IMM, Technical University of Denmark, 1996.
[4] L. G. BROWN, *A survey of image registration techniques*, ACM Computing Surveys, 24 (1992), pp. 325–376.

[5] G. E. CHRISTENSEN, *Deformable Shape Models for Anatomy*, PhD thesis, Sever Institute of Technology, Washington University, 1994.
[6] G. E. CHRISTENSEN AND H. J. JOHNSON, *Consistent image registration*, IEEE Transaction on Medical Imaging, 20 (2001), pp. 568–582.
[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), pp. 263–274.
[8] M. DROSKE AND M. RUMPF, *A variational approach to non-rigid morphological registration*, SIAM Appl. Math., 64 (2004), pp. 668–687.
[9] M. DROSKE, M. RUMPF, AND C. SCHALLER, *Non-rigid morphological registration and its practical issues*, in Proc. ICIP '03, IEEE International Conference on Image Processing, Sep. 2003, Barcelona, Spain, 2003.
[10] B. FISCHER AND J. MODERSITZKI, *Fast inversion of matrices arising in image processing*, Num. Algo., 22 (1999), pp. 1–11.
[11] ———, *Fast diffusion registration*, AMS Contemporary Mathematics, Inverse Problems, Image Analysis, and Medical Imaging, 313 (2002), pp. 117–129.
[12] ———, *Combination of automatic non-rigid and landmark based registration: the best of both worlds*, in Medical Imaging 2003: Image Processing, M. Sonka and J. Fitzpatrick, eds., Proceedings of the SPIE 5032, 2003, pp. 1037–1048.
[13] ———, *Combining landmark and intensity driven registrations*, PAMM, 3 (2003), pp. 32–35.
[14] ———, *Curvature based image registration*, J. of Mathematical Imaging and Vision, 18 (2003), pp. 81–85.
[15] ———, *A unified approach to fast image registration and a new curvature based registration technique*, Linear Algebra and its Applications, 380 (2004), pp. 107–124.
[16] J. M. FITZPATRICK, D. L. G. HILL, AND C. R. MAURER JR., *Image registration*, in Handbook of Medical Imaging, Volume 2: Medical Image Processing and Analysis, M. Sonka and J. M. Fitzpatrick, eds., SPIE, 2000, pp. 447–513.
[17] M. FORTIN AND R. GLOWINSKI, *Augmented Lagrangian Methods: Applications in the Numerical Solution of Boundary-Value Problems*, North-Holland, 1983.
[18] C. GLASBEY, *A review of image warping methods*, Journal of Applied Statistics, 25 (1998), pp. 155–171.
[19] G. GOLUB, M. HEATH, AND G. WAHBA, *Generalized cross-validation as a method for choosing a good ridge parameter*, Technometrics, 21 (1979), pp. 215–223.
[20] U. GRENANDER AND M. I. MILLER, *Computational anatomy: An emerging discipline*, March 1998.
[21] P. M. GRESHO AND R. L. SANI, *On pressure boundary conditions for the incompressible Navier-Stokes equations*, Internat. J. Numer. Methods Fluids, 7 (1987), pp. 1111–1145.
[22] E. HABER AND J. MODERSITZKI, COFIR*: Coarse and fine image registration*, Technical Report TR-2004-006-A, Department of Mathematics and Computer Science, Emory University, Atlanta GA 30322, Jun 2004.
[23] ———, *A multilevel method for image registration*, Technical Report TR-2004-005-A, Department of Mathematics and Computer Science, Emory University, Atlanta GA 30322, May 2004. Submitted to SIAM J. of Scientific Computing.
[24] ———, *Numerical methods for volume preserving image registration*, Inverse Problems, Institute of Physics Publishing, 20 (2004), pp. 1621–1638.
[25] E. HABER, J. MODERSITZKI, AND S. HELDMANN, *Computational methods for mutual information based registration*, Technical Report TR-2004-015-A, Department of Mathematics and Computer Science, Emory University, Atlanta GA 30322, Jun 2004. Submitted to Inverse Problems.
[26] W. HACKBUSCH, *Iterative Lösung großer schwachbesetzter Gleichungssysteme*, Teubner, Stuttgart, 1991. In German.
[27] J. HAJNAL, D. HAWKES, AND D. HILL, *Medical Image Registration*, CRC Press, 2001.

[28] S. HENN AND K. WITSCH, *Iterative multigrid regularization techniques for image matching*, SIAM J. on Scientific Comp., 23 (2001), pp. 1077–1093.
[29] G. HERMOSILLO, *Variational methods for multimodal image matching*, PhD thesis, Université de Nice, France, 2002.
[30] M. R. HESTENES AND E. STIEFEL, *Methods of conjugate gradients for solving linear systems*, J. Res. Nat. Bur. Stand., 49 (1952), pp. 409–436.
[31] H. J. JOHNSON AND G. E. CHRISTENSEN, *Consistent landmark and intensity-based image registration*, IEEE Transactions on Medical Imaging, 21 (2002), pp. 450–461.
[32] H. LESTER AND S. R. ARRIDGE, *A survey of hierarchical non-linear medical image registration*, Pattern Recognition, 32 (1999), pp. 129–149.
[33] E. LEVIN AND R. PIERACCINI, *Dynammic planar warping for optical character recognition*, in Proc ICASSP, vol. 3, 1992.
[34] J. B. A. MAINTZ AND M. A. VIERGEVER, *A survey of medical image registration*, Medical Image Analysis, 2 (1998), pp. 1–36.
[35] J. MODERSITZKI, *Numerical Methods for Image Registration*, Oxford University Press, 2004.
[36] J. NOCEDAL AND S. J. WRIGHT, *Numerical optimization*, Springer, New York, 1999.
[37] C. C. PAIGE AND M. A. SAUNDERS, *Solution of sparse indefinite systems of linear equations*, SIAM J. Numer. Anal., 12 (1975), pp. 617–629.
[38] A. ROCHE, *Recalage d'images médicales par inférence statistique*, PhD thesis, Université de Nice, Sophia-Antipolis, France, 2001.
[39] T. ROHLFING, C. R. MAURER, JR., D. A. BLUEMKE, AND M. 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), pp. 730–741.
[40] K. ROHR, *Landmark-based Image Analysis*, Computational Imaging and Vision, Kluwer Academic Publishers, Dordrecht, 2001.
[41] S. ROY, *Stereo without epipolar lines: A maximumflow furmulation*, International Journal of Computer Vision, 34 (1999), pp. 147–161.
[42] D. RUECKERT, L. SONODA, C. HAYES, D. HILL, M. LEACH, AND D. HAWKES, *Non-rigid registration using free-form deformations*, IEEE Transction on Medical Imaging, 18 (1999), pp. 712–721.
[43] D. SILVESTER, H. ELMAN, D. KAY, AND A. WATHEN, *Efficient preconditioning of the linearized Navier-Stokes equations*, J. Comp. Appl. Math., 128 (2001), pp. 261–279.
[44] A. W. TOGA, *Brain Warping*, Academic Press, Los Angeles, 1999.
[45] A. W. TOGA AND J. C. MAZZIOTA, *Brain mapping*, Academic Press, San Diego, 1996.
[46] A. W. TOGA AND J. C. MAZZIOTTA, *Brain Mapping: The Methods*, Academic Press, 2nd ed., 2002.
[47] U. TROTTENBERG, C. OOSTERLEE, AND A. SCHÜLLER, *Multigrid*, Academic Press, 2001.
[48] S. TUREK, *Efficient Solvers For Incompressible Flow Problems*, New York, Macmillan, 1999.
[49] S. UCHIDA AND H. SAKOE, *Piecewise linear two-dimensional warping*, in Proceedings of the 15th International Conference on Pattern Recognition, Barcelona, Spain, vol. 3, 2000, pp. 538–541.
[50] P. VIOLA AND W. M. WELLS III, *Alignment by maximization of mutual information*, (1995), pp. 16–23. IEEE 1995.
[51] G. WAHBA, *Spline Models for Observational Data*, SIAM, Philadelphia, 1990.
[52] S. K. WARFIELD, M. FERRANT, X. GALLEZ, A. NABAVI, AND F. A. JOLESZ, *Real-time biomechanical simulation of volumetric brain deformation for image guided neurosurgery*, in Proceedings of the 2000 ACM/IEEE conference on Supercomputing (CDROM), IEEE Computer Society, 2000, p. 23.
[53] B. ZITOVÁ AND J. FLUSSER, *Image registration methods: a survey*, Image and Vision Computing, 21 (2003), pp. 977–1000.