Ill-posed medicine—an introduction to image registration

Ill-posed medicine—an introduction to image registration

Bernd Fischer and Jan Modersitzki

Institute of Mathematics, University of Lübeck, Wallstraße 40, D-23560 Lübeck, Germany

E-mail: fischer@math.uni-luebeck.de and modersitzki@math.uni-luebeck.de

Abstract

Image registration is the process of aligning two or more images of the same scene taken at different times, from different viewpoints and/or by different sensors. Image registration is a crucial step in imaging problems where the valuable information is contained in more than one image. Here, spatial alignment is required to properly integrate useful information from the separate images. It is the goal of this paper to give an overview on modern techniques in this area. It turns out that the registration problem is an inverse problem which does require a sound regularization and the use of proper models. Also, the numerics have to be done with great care. We will comment on these issues and supplement it by real life examples.

1. Introduction

One of the many problems in current image processing is *image registration*, sometimes also called *fusion*, *matching* or *warping*. Very often, information obtained from multiple images appears at different poses and is of a complementary nature. Therefore spatial alignment is required to properly integrate useful information from separate images. This procedure is called *registration*. In other words, given a reference and a template image, the goal is to find a transformation such that the transformed template is similar to the reference image.

There is a large number of applications demanding registration. Areas range from art, astronomy, astro-physics, biology, chemistry, criminology, genetics, physics, or basically any area involving imaging techniques. More specific examples supported by registration are, for example, remote sensing (constructing a global picture from different partial views), security (comparing current images with a data base), robotics (tracking of objects), and in particular medicine, where computational anatomy, computer-aided diagnosis, fusion of different modalities, intervention and treatment planning, monitoring of diseases, motion correction, radiation therapy or treatment verification demand registration. Since imaging techniques, like computer tomography (CT), diffusion tensor imaging (DTI), magnetic resonance imaging

(MRI), positron emission tomography (PET), single-photon emission computer tomography (SPECT) or ultrasound (US) underwent a remarkable, fascinating and ongoing improvement in the last decade, a tremendous increase in the utilization of the various modalities in medicine takes place.

Unfortunately, no unified treatment or general theory for image registration has been established. It appears that each application area has developed its own approaches and implementations. Depending on the particular application, the focus ranges from computing time (industrial inspection, tracking), image features (remote sensing), memory (high resolution 3D images), to accuracy of a model (treatment planing). Many of the solution strategies are designed by practitioners employing various heuristics for stabilization. Care has to be taken, as image registration is ill-posed as is explained in section 3.3. Therefore, any implementation suffers from instabilities if not properly regularized.

The last decade has also witnessed several mathematical attempts to build a bridge between theory and practice. The most prominent approaches are based on stochastics, appropriate parameterizations or variational techniques. The focus of this paper is on a general mathematical model for medical image registration. This model allows one to classify most of the currently used schemes in a unified manner and thereby to discuss their pros and cons in a systematic fashion. However, it is probably impossible to provide an overview on image registration, even if one restricts oneself to the field of medical image registration. There are some books, covering parts of the field; see, e.g., [1–6]. Other important contributions aiming for an overview are, e.g., [7–17]. For important theory related contributions see, e.g. [18–21]. Among the many outstanding original publications are, e.g., [22–35].
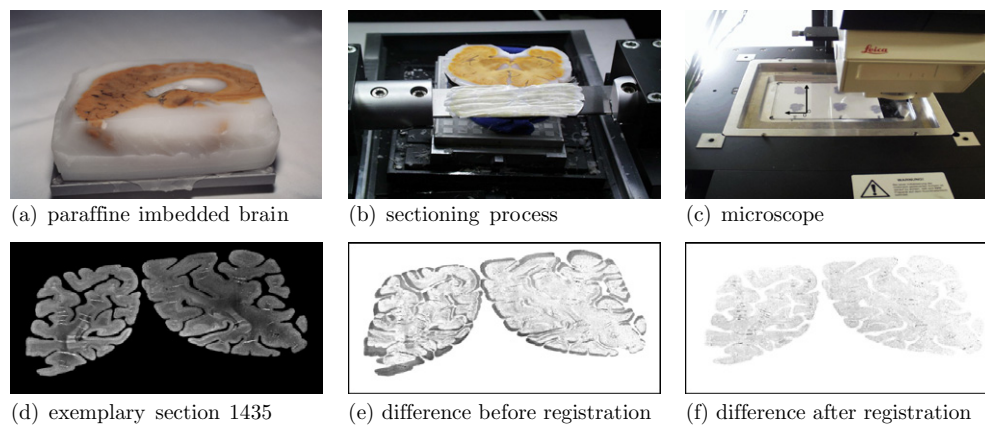
This paper is organized as follows. Some exemplary real life applications are presented which are intended to give an idea on the widespread medical registration problems as well as the application-driven demands on a solution. In section 3, the general model is presented gradually and its main ingredients are discussed and commented. Finally, section 4 touches upon the state-of-the-art numerics for solving the challenging registration problems. We close with some remarks on future challenges.

2. Motivation

To illustrate the basic challenges and the diversity of medical image registration, three exemplary applications are presented. The first one is concerned with the alignment of images from serial sectioning of a human brain (image courtesy Oliver Schmitt), the second example deals with the bending of a human knee (a challenging test case supplied by Philips Research, Hamburg), whereas the third one is a typical multimodality example (images from FUSION, Germany).

2.1. Example: serial sectioning

Especially in anatomy, serial sectioning is a common technique for exploring tissue properties down to a cellular level. Figure 1 illustrates a typical work flow. Here, an organ (the brain of a human donator) is prepared and embedded into paraffine wax (a). Using a sliding microtome, the tissue is sectioned into thin slices (b), which amount to about 6.000 for a whole human brain. The slices are stained, mounted and covered by an objective slide. These slides serve as the basis for a microscopic study (c). The overall process destroys the three-dimensional correspondences of tissue: the fixation of the tissue on the objective slide is almost arbitrary and in addition the sectioning procedure itself introduces nonlinear distortions of the tissue.

(a) paraffine imbedded brain   (b) sectioning process   (c) microscope

(d) exemplary section 1435   (e) difference before registration   (f) difference after registration

Figure 1. Histological serial sectioning (image courtesy Oliver Schmitt, Institute of Anatomy, University of Rostock, Germany): paraffine embed human brain (a), slicing microtome with tissue section on the blade (b), microscope exploring objective slide (c), digitized section 1435 (d), difference with respect to the consecutive section before (e) and after registration (f).

(a) detail with grid   (b) 3D reconstruction
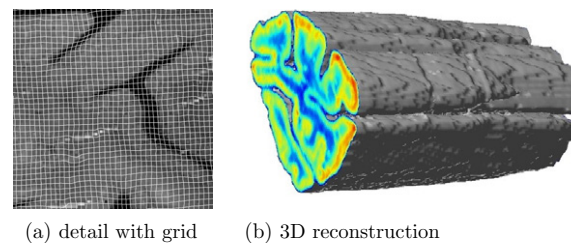
Figure 2. Detail of a transformed section with a visualization of the grid (a) and a stack of registered section (b) (100 sections, left hemisphere, zoom in the *z*-direction).
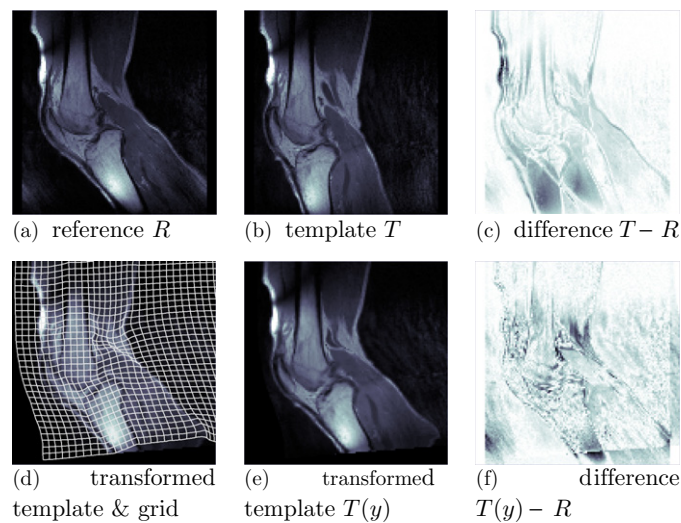
The rearrangement of the original tissue correspondence constitutes a typical registration problem.

Figure 1 also shows an exemplary section (d) and the difference in the consecutive section before (e) and after (f) registration. As it is apparent, all structural differences have been removed by registration. Figure 2 shows a stack of sections after the registration (b). To appreciate the computed transformation, a detail of a transformed grid is displayed for a sample slice (a). The numerical details are provided in section 3.7.

2.2. Example: human knee

Next, an alignment of two MR scans of a human knee is presented (data courtesy by Thomas Netsch, Philips Research, Hamburg, Germany). Though this example has no clinical relevance, it shows many of the features and challenges of typical registration problems but at the same time allows for an intuitive visualization and interpretation.

To keep the issue of interest simple, we display 2D cross sections from 3D MRI's of a human knee. Figure 3 shows the images to be registered (a, b). Assuming that these two images are taken before and after an intervention, the objective is to automatically detect correspondences between points in both images. Due to a bending of the knee, these correspondences are non-trivial, as can be observed from a visual inspection or from the

(a) reference R (b) template T (c) difference $T - R$

(d) transformed template & grid   (e) transformed template $T(y)$ (f) difference $T(y) - R$

Figure 3. Local rigidity constrained registration of a human knee (data courtesy of Thomas Netsch, Philips Research, Hamburg, Germany): the nonlinear transformation has been kept locally rigid to preserve bone structures.

difference image (c). Figure 3 also illustrates the transformation (visualized by a transformed grid (d)), the transformed image (e) and the distance after a registration (f). The registration, which is discussed in section 3.7, enables a direct comparison of the images.

2.3. Example: image-guided surgery

The next example discusses another challenging real life application for image registration, namely the use of intraoperative imaging to facilitate an image-guided surgery. This work is part of the FUSION project (Future Environment for Gentle Liver Surgery Using Image-Guided Planning and Intra-Operative Navigation) granted by the BMBF (German Federal Ministry of Education and Research); see www.somit-fusion.de.

Image-guided surgery aims to provide the surgeon with accurate in situ spatial information regarding the location of pathology during intervention. FUSION focuses on liver resection and ablation. Typically the removal of tumors leaves the surgeon with a dilemma: remove all tumor tissues while sparing a sufficient portion of the patient's liver. Preoperative imaging is an important tool for a qualified planning. Figure 4 displays a typical plan for a liver resection (planing by MeVis, Bremen, Germany).

Due to liver movement during surgery, registration is required to match the planning data to the current situation. In order to get a real-time measurement of the current configuration a fast 2D freehand US image acquisition is used which is optically tracked. A US compounding is used to generate a 3D US volume.

For illustration purposes, a simplified 2D problem is considered. Figure 4 depicts a cross section of the preoperative planning MR data (a) and a corresponding US view (c). Apart from the fact that one faces the problem of only partial overlap of the US image and the MR data one has also to deal with different modalities. This first problem is bypassed in the FUSION project by an initial pre-registration of the region of interest (ROI) using some user-supplied landmarks and the second problem by using NGF [36], a particular modality-
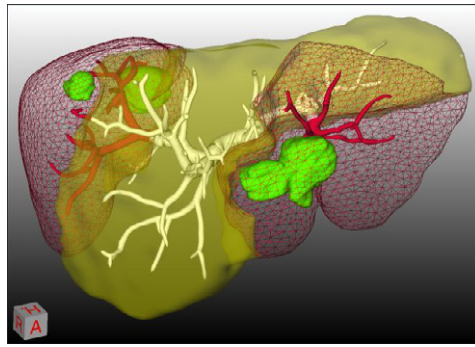
Figure 4. Liver resection planning by MeVis, Bremen, Germany.

(a) MR template   (b) MR ROI   (c) US reference   (d) transformed MR ROI
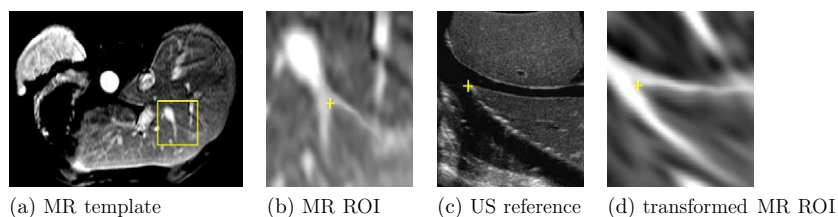
Figure 5. MR/US registration of a human liver (data by FUSION): cross section of the MR data (a), region of interest (b), US reference (c) and detail of the transformed ROI (d).

independent distance measure. The overall result is shown in figure 5(d), the numerical details are provided in section 3.7.

3. Problems and solutions

Medical image registration algorithms usually estimate the transformation by either following a data-driven flow or minimizing a certain cost function. Flow approaches are very similar to approaches in so-called optical flow in computer vision, a technique used to estimate motion in an image sequence like video. In image registration, flow approaches have been studied for example in [24, 37, 35]. Here, the focus is on registration via optimization. In particular, a variational-based approach is discussed. This approach not only allows for a sound mathematical treatment but also enables a characterization and classification of some of the most prominent methodologies.

Optimization-based registration can be classified according to the space to which the wanted deformation belongs: rigid or affine linear registration algorithms depend on only a few parameters, while spline-based approaches may use a very high-dimensional transformations space. Rigid registration methods are in general very limited and not descriptive enough and therefore nonrigid approaches are to be used in most applications, like those presented in this paper. More generally, the search space might be explicitly given (e.g. spanned by splines) or implicitly determined by a variational formulation (e.g. using Sobolev semi/norm regularization).

As is outlined in section 3.3, image registration is ill-posed which results in ill-conditioning, instability of solutions and a highly non-convex cost function. Regularization is

introduced to alleviate these issues and also to incorporate user knowledge into the problem formulation. Therefore, the image registration is usually phrased as an optimization problem, where the cost function consists of a similarity measure, a regularization and additional penalty terms that discourage undesirable transformations or constraints that rule out unfeasible solutions.

Two major directions can be distinguished. One direction is based on a fixed Tichonov regularization, penalizing against a fixed starting point, for example a pre-registration, while the other direction is based on iterative Tichonov regularization where the update is regularized. The second approach generally allows for more flexible transformations and thus is an interesting tool for inter-patient registration. The first approach is more restrictive and, with a proper choice of the regularization parameter, can be used to ensure a one-to-one transformation, which is very valuable for some applications.

In the following, it is assumed that the images $\mathcal{T}$ and $\mathcal{R}$ are given, smooth, and compactly supported functions on a rectangular domain $\Omega \subset \mathbb{R}^d$, where d denotes the data dimensionality. Since the given data are typically discrete, an interpolation or approximation step might become necessary, in particular for the computation of the transformed template $\mathcal{T}[y]$ (cf section 3.1). Using an appropriate distance measure $\mathcal{D}$ (cf section 3.2) and a regularizer $\mathcal{S}$ (cf section 3.3), the wanted transformation is a solution of the optimization problem

$$\mathcal{D}[T[y], \mathcal{R}] + \alpha \mathcal{S}[y] + \beta \mathcal{P}[y] = \min \qquad \text{subject to} \quad y \in \mathcal{M}, \tag{1}$$

where α is a regularization parameter, $\mathcal{P}$ is a penalty or soft constrained (cf section 3.4), β is a penalty parameter and $\mathcal{M}$ is a set of admissible transformations (cf section 3.5). Note that a functional setting is used. A major advantage of this continuous formulation is that the following steps are consistent and justifiable. Dealing with discrete data requires interpretation and work around at later stages. For example, already a rotation of an object is a non-trivial operation within a discrete setting.

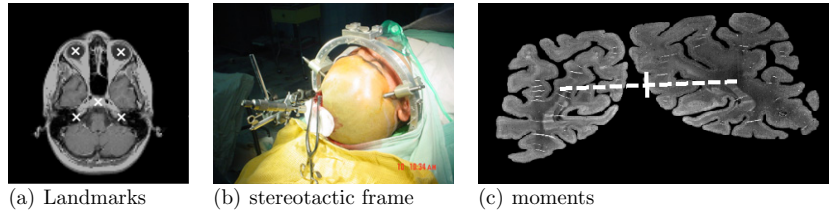### 3.1. Transforming images, the forward problem

A major ingredient that is often overlooked and underestimated is the computation of the transformed image $\mathcal{T}[y]$. A first step is the interpolation of the typically discrete given data T_i. In particular, the d-linear interpolation is widely used. The advantages of d-linear interpolation are a continuous representation, a min/max principle ($\min\{T_i\} \leqslant \mathcal{T}(x) \leqslant \max\{T_i\}$) and its fast computation time. A drawback is a resulting non-differentiable overall objective function. As a consequence, only rather slow optimization techniques can be applied. Higher order spline interpolation is a commonly used alternative, resulting in a smoother, differentiable image model. The drawbacks of spline interpolation are under-/overshooting (Gibb's phenomena) and its slightly higher computation time [32, 38–41].

Sometimes, in addition, approximation or scale-space ideas are used for convexification of the objective function. Rather than interpolating the measurements that are usually corrupted by noise, a smoother approximation is used [41, 42]. Thus image details are omitted and the danger of being trapped into local minima can be reduced. Figure 6 illustrates these concepts, and a spline interpolation (a), a smoother approximation (b) and a very sooth approximation (c). Note that the images are represented with respect to different scales rather than at a different resolution, which will be discussed in section 3.3.

Based on the continuous representation of $\mathcal{T}$, the above so-called Eulerian framework is obtruding, $\mathcal{T}[y](x) := \mathcal{T}(y(x))$. An alternative, though less commonly used approach is the so-called Lagrangian map, where basically $\mathcal{T}^{\text{Lagrange}}[y](x) := \mathcal{T}(y^{-1}(x))$ (in a proper formulation, the inverse transformation is not required).

(a) fine   (b) coarse   (c) ultra coarse

(d) fine   (e) coarse   (f) ultra coarse
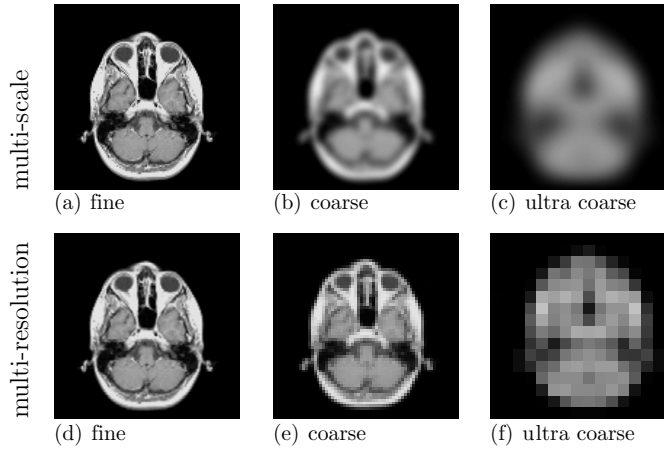
Figure 6. Multi-scale image approximation.

(a) Landmarks   (b) stereotactic frame   (c) moments

Figure 7. (a) MR scan of a human head with manually set landmarks; (b) stereotactic frame; images from [43], see http://cns.bu.edu/~rash/paper.htm; (c) image with principal axis deduced from second-order moments.

3.2. Distance measures

The next step is to quantify image similarity, proximity or distance. Several approaches have been proposed. These measures are based either on image features (like for example markers—tags that are attached to objects before imaging, landmarks—tags deduced after imaging, moments—statistical quantities), substructures (surfaces, level sets, crest lines) or volumetric data. Here, landmark and L_2-norm-based distance measures are discussed exemplarily.

For landmark-based approaches, it is assumed that a number of points $r^1, \ldots, r^m \in \mathbb{R}^d$ are specified in the reference image and the corresponding points $t^1, \ldots, t^m$ are located in the template image (see figure 7) and the MR/US example. A commonly used distance measure then reads

$$\mathcal{D}^{\mathrm{LM}}[y] = \sum \|y(r^j) - t^j\|^2_{\mathbb{R}^d}, \tag{2}$$

where extensions using different norms are used as well (see, e.g., [44]).

Several approaches are employed to compute the wanted transformation. The first one is to restrict the search space to a typically low-dimensional parametric space $\mathcal{M}$. Common choices for $\mathcal{M}$ are based on rigid, affine linear or spline transformations. Clearly, the transformation is a linear combination of some basis functions q_k, i.e. $y(x) = \sum q_k(x)w_k$, and it is computed by solving a least-squares problem for the coefficients:

$$\text{minimize } \mathcal{D}^{\mathrm{LM}}[y] \qquad \text{subject to} \quad y \in \mathcal{M}. \tag{3}$$

A second approach is to optimize a certain smoothness $\mathcal{S}$ of the transformation, i.e. minimize $\mathcal{S}[y]$ subject to $\mathcal{D}^{\text{LM}}[y] = 0$. This more general concept is discussed further in section 3.3. Particularly for landmark-based registration, a thin-plate-spline bending energy is used as a smoothness measure and the resulting registration is called the thin-plate spline registration (see also [23, 44–46]). Following representer theory [46], it can be shown that the solution to this problem belongs to a certain parameterizable space $\mathcal{M}^{\text{TPS}}$, spanned essentially by translates of a radial basis function related to $\mathcal{S}$. Hence, this approach is a special case of (3). Numerical solutions can be obtained by solving a linear system of equations for the coefficients.

In practice, in particular in the presence of noise, the exact location of landmarks is a tricky problem. Therefore, the interpolation condition is sometimes replaced by a weighted approximation condition

$$\text{minimize } \mathcal{D}[y] + \theta\mathcal{S}[y] \qquad \text{subject to} \quad y \in \mathcal{M}^{\text{TPS}}.$$

Setting $\theta = 0$ returns (3) whereas $\theta = \infty$ results in a very smooth transformation which may not fulfil the interpolation constraints; see [4, 44] for a more detailed discussion.

Advantages of landmark-based registrations are that their interpretation is easy and intuitive, solutions can be computed fast and efficiently. However, neither the determination of landmarks nor the identification of proper locations of corresponding landmarks is an easy task and a scheme for a fully automatic detection of landmarks in medical images is still missing.

A striking property of landmark-based registration is its build in dimension reduction. The images are replaced by just a few outstanding points. Alternative approaches along these lines consider 3D space curves [47], the surface of the given objects [48, 49] or other geometric features [50]. Sometimes, these methods are additionally intermixed with the use of a proper biomechanical model [51, 52]. However, all these methods benefit and at the same time suffer from the fact that the algorithm only sees parts of the images and may produce unwanted results for the remaining parts.
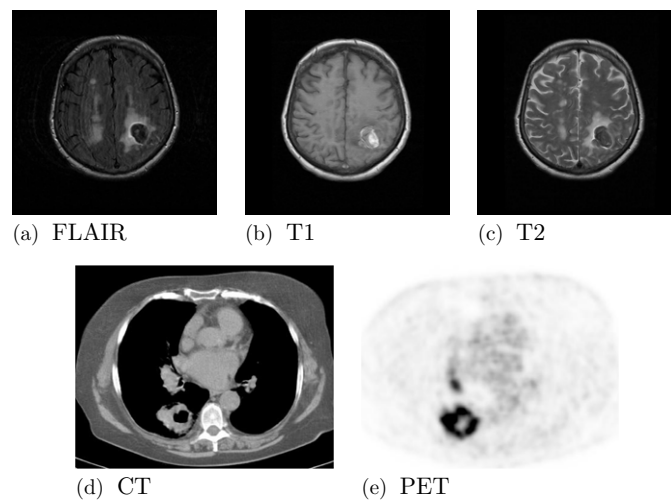
To overcome the just-mentioned drawbacks one is bound to consider the whole images. Probably the most intuitive volumetric distance measure is the so-called sum of square difference (SSD) (also L_2-norm of the image difference or energy of image distance),

$$\mathcal{D}^{\text{SSD}}[y] = \|\mathcal{T}[y] - \mathcal{R}\|^2_{L_2(\Omega)} = \int_\Omega (\mathcal{T}[y] - \mathcal{R})^2 \, dx. \tag{4}$$

It has been proven to be robust and very effective for images of one modality, e.g., the serial sectioning or the human knee. It is based on the assumption that $\mathcal{T}[y](x) = \mathcal{R}(x)$, which most often does not hold. For image modalities like CT, one has a normalized protocol; however, for MRI this normalization does not hold (T1, T2, FLAIR, 1.5, 3, 7 Tesla; see figures 8(a), (b) and (c) for an example). Moreover, applications of particular interest are the registration of images originating from different imaging devices, like MR and US as in the previous example or like CT (displaying anatomy) and PET (displaying functionality); see figures 8(d) and (e) for an example.

For images of different modalities, cross-correlation (and variants), normalized gradient fields [36] and in particular mutual information [25, 26] are commonly used alternatives; see, e.g. [4, 53, 54] for a detailed overview. Based on either a histogram or the Parzen–Window-based estimator $q_{\mathcal{T},\mathcal{R}}$ for the joint density of the intensity distribution of the images $\mathcal{T}$, $\mathcal{R}$, mutual information is a normalized entropy:

$$\mathcal{D}^{\text{MI}}[y] = H(q_{\mathcal{T}[y]}) + H(q_{\mathcal{R}}) - H(q_{\mathcal{T}[y],\mathcal{R}}),$$

$$H(q) = -\int_{\mathbb{R}^2} q(\tau, \rho) \log \rho(\tau, \rho) d(\tau, \rho),$$

(a) FLAIR   (b) T1   (c) T2

(d) CT   (e) PET

Figure 8. MR images of a brain tumor ((a), (b) and (c)) acquired with different protocols (images are taken from CASIMAGE http://www.pubimage.hcuge.ch); CT (d) and PET (e) images of a lung tumor (images courtesy by the Institute of Nuclear Medicine, University Hospital, Zürich).

(a) Reference   (b) template
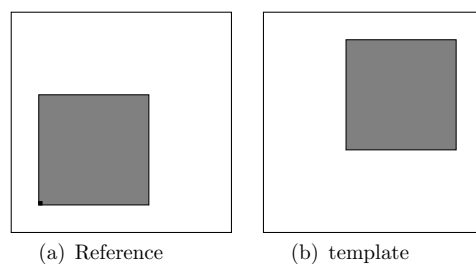
Figure 9. Ambiguous results for image registration.

where $q_{\mathcal{T}[y]}$ and $q_{\mathcal{R}}$ denote the marginal densities, respectively. Mutual information seems to be the most powerful distance measure, but as a consequence may cause trouble for the optimization routine; see, e.g, [36, 55, 56].

3.3. Regularization

As already pointed out several times, image registration is inherently ill-posed. For every spatial location $x \in \Omega \subset \mathbb{R}^d$, one asks for a vector valued transformation $y(x) \in \mathbb{R}^d$, but in general only scalar information (image intensity) is provided. Regularization is thus important and inevitable. A more subtle point is emphasized in figure 9, showing a gray rectangle on a white background. A possible registration is a translation, transforming for example the top-left corner of the template onto the top-left corner of the reference. However, rotating the image about 180° around the center of the domain results in an alternative meaningful, though completely different, solution. Note that both solutions are rigid transformations. Finally, if one does allow for more complex transformations, the transformation is indefinite in areas of constant gray values. In fact, any permutation of points of equal color yields the same distance. In the extreme case when both images are constant, any transformation is a solution.

9

The problem becomes even more delicate in the presence of noise. Adding one black dot to either of the four corners of both images would lead to a unique rigid solution (which is then completely determined by noise). Fortunately, images with more structure (more variations in intensities) are in general less vulnerable.

The focus of regularization in the image registration is on the existence rather than on the uniqueness of the underlying optimization problem (1). To ensure uniqueness alternative strategies (penalties, constraints), as discussed later on, are considered. Additionally, the choice of an appropriate starting guess is a crucial ingredient of any registration scheme. In FUSION, a landmark-based pre-registration is used to provide an improved starting guess. Regularization is achieved by selecting an admissible set $\mathcal{M}$ and/or by choosing an application-dependent Sobolev semi-norm (bi-linear form),

$$\mathcal{S}[y] = \|\mathcal{B}[y]\|^2_{L_2(\Omega)^q} = \int_\Omega \langle \mathcal{B}[y], \mathcal{B}[y] \rangle \, dx, \tag{5}$$

where in particular the following choices are common (omitting some parameters):

$$\mathcal{B}^{\mathrm{diff}}[y] = (\nabla y^1, \ldots, \nabla y^d) \qquad \text{(diffusion)},$$
$$\mathcal{B}^{\mathrm{elas}}[y] = (\nabla y^1, \ldots, \nabla y^d, \nabla \cdot y) \qquad \text{(elastic)}, \tag{6}$$

$$\mathcal{B}^{\mathrm{curv}}[y] = (\Delta y^1, \ldots, \Delta y^d) \qquad \text{(curvature)}, \tag{7}$$

see [4, 22, 57–59]. Typically, the regularization is used for the displacement u, where $y(x) = x + u(x)$; see, e.g. [4] for a detailed discussion. Unlike other ill-posed problems, the particular regularizer and in particular the choice of boundary conditions (not discussed here) can effect the solution considerably. As pointed out for the above example, the influence becomes less prominent, if the images have more structure. For example, treating the serial sections with different regularizers yields visually undistinguishable results.

Another important issue is the choice of the regularization parameter. Typically, this parameter is hand-tuned. For an automatic choice of the regularization parameter, a continuation method has been suggested in [60].

3.4. Soft constraints, penalties

Another step toward a more reliable outcome of the registration is to add a penalty $\mathcal{P}$ (see (1)), penalizing unwanted solutions. In contrast to a regularizer, which is needed to guarantee solutions of the registration problem, the penalty is an add-on which may serve as an instrument for incorporating user knowledge.

Examples for penalty terms include the deviation from user-supplied landmarks [61] (i.e. setting $\mathcal{P}^{\mathrm{LM}}[y] = \mathcal{D}^{\mathrm{LM}}[y]$) volume preservation [62] ($\mathcal{P}^{\mathrm{VP}}[y] = \int_\Omega \log(\det \nabla y)^2 \, dx$), or local rigidity [63–66]. In addition to anatomical issues (branching points of vessel should be co-registered), the landmark penalty not only provides a specific starting guess but also pushes the scheme toward a particular solution. Consequently, as a by-product, this penalty helps to avoid unwanted local minima. In the example of the square (figure 9), three additional landmarks would lead to a perfect solution, if the correspondence of landmarks is correctly established. Volume preservation is a more application-oriented penalty. Any unconstrained registration scheme most likely would lead to a change of volume if the objects in question are of different size; otherwise it would make a poor job. However, a registration-induced change of volume may lead to false diagnoses with severe consequences. For example, the evaluation of tumor growth in follow-up screening might be effected by an artificial volume

change due to registration. The idea of local rigidity is to transform hard tissue like bones rigidly but still allow for a globally nonrigid transformation elsewhere; see the knee example. To this end, rigidity is considered on particular segments (sub-domains) by penalizing deviation from linearity, orthogonality and orientation preservation.

3.5. Constraints

In contrast to the penalty approach, where unwanted solutions are penalized, hard constraints are used to rule out these transformations. The feasible set is either described explicitly or in terms of constraints $\mathcal{C}$ (see (1)). Explicit descriptions are, for example, parameterized transformations like rigid, affine linear or diffeomorphic spline transformations [67–69]. Implicit constraints range from equality constraints $\mathcal{C}[y] = 0$ like landmarks constraints [70]

$$\mathcal{C}^{\text{LM}}[y] = (y(r^j) - t^j)_{j=1}^m,$$

space segregation [71], volume preservation [41]

$$\mathcal{C}^{\text{VP}}[y](x) = \det \nabla y(x) - 1, \qquad x \in \Omega,$$

or local rigidity [72], to inequality constraints like

$$k(x) \leqslant \mathcal{C}^{\text{VC}}[y](x) = \det \nabla y(x) \leqslant K(x), \qquad 0 < \kappa(x) \leqslant K(x),$$

which ensure that the transformation is one-to-one, see also [73]. The displayed equality constraints match the penalty terms of the previous section. However, instead of penalizing unwanted solutions, the constraints completely rule out them. In particular for volume preservation this is of great value. The penalty term leads to a small integral, which still allows for a considerable volume change in small regions like tumors. In contrast, the volume constraint ensures volume preservation in any part of the image.

3.6. Optimization

In this section, we briefly comment on the numerical solution of the outlined registration problems. The backbone of the solution schemes is optimization techniques. For a general overview on numerical optimization techniques, see, e.g., [74–77].

In the literature one may find essentially three different optimization approaches within a registration procedure. As already pointed out, often the registration formulation is based on practitioners heuristics. Based on the rough modeling, the objective functions are in general not differentiable or even not continuous. Thus, the optimization techniques are in general limited to basic schemes like Nelder–Mead's simplex method or Powell's downhill method. In addition particular implementations may act as regularizer and stabilize the overall scheme.

The second class of approaches aims for the necessary condition for a minimizer of the objective function in (1), which turns out to be a nonlinear system of partial differential equations

$$f[T, R, y](x) + \alpha \mathcal{A} y(x) = 0,$$

where the so-called force field f is related to the particular distance measure and the partial differential operator $\mathcal{A}$ corresponds to the regularizer. Subsequently, the partial differential equation, equipped with appropriate boundary conditions, is solved by some sort of discretization technique. Various techniques have been proposed; for an overview, see [4] or the original papers [24, 30, 37, 57–59, 78, 79]. The overall scheme is also known under the notion *optimize–discretize*, which indicates the order of its major ingredients.

A more recent approach is known as the *discretize–optimize* technique. It is based on a sequence of nested discretizations of (1) [60]. Each discretization leads to a finite-dimensional optimization problem. Starting with a coarse discretization, a minimizer is computed which then serves as a starting point for the optimization problem associated with the finer discretization. If the coarse grid solution is a good starting guess, the optimization on the finer discretization is more like a correction step and is expected to be performed with low computational costs. To arrive at a fast converging scheme, care has to be taken during modeling. If the objective function is sufficiently smooth, even Newton-type optimization techniques can be applied. A major advantage, as compared to the PDE-based approaches, is that sound stopping rules and line search procedures are available. Focusing on optimization aspects rather than smoothing and grid transfer operations makes a difference to similar multigrid-type techniques; see, e.g., [80, 81].

No matter which strategy is employed, there is always the danger of being trapped by an unwanted minimum. To reduce this problem, four main ideas are given, which may be seen as convexification of the overall optimization problem. The first is to start with a pre-registration that generates an acceptable starting guess in a neighborhood of a wanted minimizer. The second step is to add a regularization $\mathcal{S}$ and/or penalties $\mathcal{P}$, penalizing unwanted solutions. The third idea is to solve the problem in a multilayered fashion, i.e. in a sequence of discretizations ranging from coarse to fine, where on a coarse level only main characteristics of the underlying images are visible and resolved and where subsequently on finer levels more and more characteristics are added and resolved. Here, the point is to use a multi-resolution of the images. The fourth concept is along the same lines, details are suppressed for the start-up phase and then are gradually added. This way one arrives at a multi-scale technique which of course can be mixed with the multi-resolution approach.

3.7. Numerical details of the examples

In this section we summarize the particular building blocks used for the motivating examples in section 2. All results were obtained by solving problem (1) with a multi-resolution approach running from coarse to fine. On every level, a quasi-Newton scheme serves as a working horse for the optimization, where the Hessian of the data fitting term is approximated by first-order derivatives.

3.7.1. Serial sectioning. For this histological serial sectioning problem, the L_2-norm distance measure (4) and the elastic potential regularizer (6) are exploited. This choice is motivated by the fact that the tissue is homogeneous paraffine wax and is supposed to deform elastically [82]. A pre-processing step has been used to homogenize the intensity values.

3.7.2. Human knee. For the human knee, again, the L_2-norm (4) and the elastic potential (6) are used as main building blocks. In addition, the transformation is constrained to be rigid on the tibia and femur parts:

$$y|_{\text{tibia}}, y|_{\text{femur}} \in \{y : \mathbb{R}^d \to \mathbb{R}^d | y(x) = Ax + b, A^\top A = I_d, \det A = 1\}.$$

To do so, a segmentation of these areas has to be supplemented. A Lagrangian framework for the problem is used to keep the constraints linear. The overall implementation is based on standard SQP techniques.

3.7.3. Image-guided surgery. For the registration of multi-model liver data the normalized gradient field [36] and the curvature regularizer (7) have been used. The registration is constrained by the correspondence of user-supplied landmarks $r^j, t^j, j = 1, \ldots, m$:

$$y \in \{y : \mathbb{R}^d \to \mathbb{R}^d | y(r^j) = t^j, j = 1, \ldots, m\}.$$

In addition, the multiscale approach has been envoked in order to smooth the noisy ultrasound image. For the optimization, a reduced basis of the manifold is used.

4. Summary, conclusions, outlook

The paper introduces medical image registration, shows characteristic examples and explains state-of-the-art techniques. Based on a general variational modeling, substantial building blocks and problems are considered and are discussed. In particular, the inherently ill-posedness has been addressed and several techniques to overcome this problem have been outlined. Common numerical approaches have been introduced briefly.

It has been shown that registration is an important task in various disciplines. In particular, medical applications have been discussed which are likely to become even more important as imaging techniques are undergoing an amazing development. As of today, many of the practical problems have been addressed, though important issues are still to be solved, resulting in a very active research community.

Probably the most important open issue is the objective validation of registration results. Todays validation approaches are mostly based on subjective eye-ball norms of trained experts, on artificial examples or phantoms, or on tagged data, which are of limited significance in a clinical environment (for promising approaches toward validation see also http://www.insight-journal.org/RIRE or http://www.NIREP.org). Another issue is the integration of user knowledge, which might be seen as a step toward well-posedness. Recent approaches based on constrained registration proved to be very promising. A further bottleneck is computing time which still limits the applicability of many techniques. Current scanners can produce 1024^3 data resulting in systems with approximately 3×10^9 unknowns which have to be processed preferably in real time. Finally, most schemes depend on a variety of parameters which most often have to be hand-tuned. An automatization of parameter and regularization choices is another important research area.

New trends include the integration of several imaging tasks into a unified framework. Examples include the combination of segmentation and registration [83] or an automatic correction of gray value disturbances as they can appear for example in serial sectioning [84].

References

[1] Highnam R and Brady M 1999 *Mammographic Image Analysis (Kluwer Series on Medical Image Understanding)* (Dordrecht: Kluwer)
[2] Faugeras O, Aubert G and Kornprobst P 2002 *Mathematical Problems in Image Processing* (New York: Springer)
[3] Yoo T S 2004 *Insight into Images: Principles and Practice for Segmentation, Registration, and Image Analysis* (Wellesey, MA: A K Peters)
[4] Modersitzki J 2004 *Numerical Methods for Image Registration* (Oxford: Oxford University Press)
[5] Goshtasby A A 2005 *2D and 3D Image Registration* (New York: Wiley)
[6] Scherzer O 2006 *Mathematical Models for Registration and Applications to Medical Imaging* (New York: Springer)
[7] Bajcsy R and Kovačič S 1989 Multiresolution elastic matching *Comput. Vis. Graph. Image Process.* **46** 1–21
[8] Brown L G 1992 A survey of image registration techniques *ACM Comput. Surv.* **24** 325–76

[9] van den Elsen P A, Pol E-J D and Viergever M A 1993 Medical image matching—a review with classification *IEEE Eng. Med. Biol.* **12** 26–38

[10] Maintz J B A and Viergever M A 1998 A survey of medical image registration *Med. Image Anal.* **2** 1–36

[11] Glasbey C 1998 A review of image warping methods *J. Appl. Stat.* **25** 155–71

[12] Lester H and Arridge S R 1999 A survey of hierarchical non-linear medical image registration *Pattern Recognit.* **32** 129–49

[13] Fitzpatrick J M, Hill D L G and Maurer C R 2000 Image registration *Handbook of Medical Imaging, Volume 2: Medical Image Processing and Analysis* ed M Sonka and J M Fitzpatrick (Bellingham, WA: SPIE) pp 447–513

[14] Hajnal J, Hawkes D and Hill D 2001 *Medical Image Registration* (Boca Raton, FL: CRC Press)

[15] Hill D L G, Batchelor P G, Holden M and Hawkes D J 2001 Medical image registration *Phys. Med. Biol.* **46** R1–45

[16] Toga A and Thompson P 2001 The role of image registration in brain mapping *Image Vis. Comput.* **19** 3–24

[17] Zitová B and Flusser J 2003 Image registration methods: a survey *Image Vis. Comput.* **21** 977–1000

[18] Dupuis P, Grenander U and Miller M I 1998 Variational problems on flows of diffeomorphisms for image matching. *Q. Appl. Math.* **56** 587–600

[19] Clarenz U, Droske M and Rumpf M 2002 Towards fast non-rigid registration *Inverse Problems, Image Analysis and Medical Imaging, AMS Special Session Interaction of Inverse Problems and Image Analysis* vol 313 (Providence, RI: American Mathematical Society) pp 67–84

[20] Hinterberger W, Scherzer O, Schnörr C and Weickert J 2002 Analysis of optical flow models in the framework of calculus of variations *Num. Funct. Anal. Opt.* **23** 69–82

[21] Droske M and Rumpf M 2004 A variational approach to non-rigid morphological registration *SIAM Appl. Math.* **64** 668–87

[22] Broit C 1981 Optimal registration of deformed images *PhD Thesis* Computer and Information Science, University of Pensylvania

[23] Bookstein F L 1989 Principal warps: thin-plate splines and the decomposition of deformations *IEEE Trans. Pattern Anal. Mach. Intell.* **11** 567–85

[24] Christensen G E 1994 Deformable shape models for anatomy *PhD Thesis* Sever Institute of Technology, Washington University

[25] Collignon A, Vandermeulen A, Suetens P and Marchal G 1995 3D multi-modality medical image registration based on information theory *Comput. Imag. Vis.* **3** 263–74

[26] Viola P A 1995 Alignment by maximization of mutual information *PhD Thesis* Massachusetts Institute of Technology

[27] Studholme C, Hill D L G and Hawkes D J 1996 Automated 3d registration of MR and CT images of the head *Med. Image Anal.* **1** 163–75

[28] Rueckert D, Hayes C, Studholme C, Summers P, Leach M and Hawkes D J 1998 Non-rigid registration of breast MR images using mutual information *MICCAI 98: Med. Image Comput. and Computer-Assistant Intervention* ed A C F Colchester, S L Delp and W M Wells III (Berlin: Springer) pp 1144–52

[29] Trouvé A 1998 Diffeomorphisms groups and pattern matching in image analysis *Int. J. Comput. Vis.* **28** 213–21

[30] Thirion J P 1998 Image matching as a diffusion process: an analogy with Maxwell's demons *Med. Image Anal.* **2** 243–60

[31] Rueckert D, Sonoda L, Hayes C, Hill D, Leach M and Hawkes D 1999 Non-rigid registration using free-form deformations *IEEE Trans. Med. Imag.* **18** 712–21

[32] Kybic J and Unser M 2000 Multidimensional elastic registration of images using splines *ICIP in Vancouver (10–13.09.2000)* vol 2, pp 455–8

[33] Christensen G E and Johnson H J 2001 Consistent image registration *IEEE Trans. Med. Imag.* **20** 568–82

[34] Twining C J and Marsland S 2004 Constructing diffeomorphic representations for the groupwise analysis of non-rigid registrations of medical images *IEEE Trans. Med. Imag.* **23** 1006–20

[35] Keeling S L and Ring W 2005 Medical image registration and interpolation by optical flow with maximal rigidity *J. Math. Imag. Vis.* **23** 47–65

[36] Haber E and Modersitzki J 2007 Intensity gradient based registration and fusion of multi-modal images *Methods Inf. Med.* **46** 292–9

[37] Bro-Nielsen M 1996 Medical image registration and surgery simulation *PhD Thesis* IMM, Technical University of Denmark

[38] Unser M, Aldroubi A and Eden M 1993 B-spline signal processing: part I—theory *IEEE Trans. Signal Process.* **41** 834–46

[39] Unser M l, Aldroubi A and Eden M 1993 B-spline signal processing: part II—efficient design and applications *IEEE Trans. Signal Process.* **41** 821–32

[40] Thévenaz P, Blu T and Unser M 2000 Image interpolation and resampling *Handbook of Medical Imaging, Processing and Analysis* ed I N Bankman (San Diego, CA: Academic) pp 393–420

[41] Haber E and Modersitzki J 2004 Numerical methods for volume preserving image registration *Inverse Problems* **20** 1621–38

[42] Haber E and Modersitzki J 2005 A scale space method for volume preserving image registration *Proc. 5th Int. Conf. on Scale Space and PDE Methods in Computer Vision* pp 1–8

[43] Abbassian A H, Shahzadi S, Afraz S R, Fazl A and Moradi F 2001 Tactile discrimination task not disturbed by thalamic stimulation *Stereotact. Funct. Neurosurg.* **76** 19–28

[44] Rohr K 2001 Landmark-based image analysis *Computational Imaging and Vision* (Dordrecht: Kluwer)

[45] Duchon J 1976 Interpolation des fonctions de deux variables suivant le principle de la flexion des plaques minces *RAIRO Anal. Numér.* **10** 5–12

[46] Light W A 1996 Variational methods for interpolation, particularly by radial basis functions in *Numerical Analysis 1995* (ed D F Griffiths G A Watson) (London: Longmans) pp 94–106

[47] Guéziec A and Ayache N 1994 Smoothing and matching of 3-d space curves *Int. J. Comput. Vis.* **12** 79–104

[48] Pellizari C A, Chen G T Y, Spelbring D R, Weichsellbaum R R and Chin C T 1989 Accurate three dimensional registration of CT, PET, and/or MR images of the brain *J. Comput. Assist. Tomogr.* **30** 20–6

[49] Besl P J and McKay N D 1992 A method for registration of 3-d shapes *IEEE Trans. Pattern Anal. Mach. Intell.* **14** 239–56

[50] Hahn D, Wolz G, Sun Y, Hornegger J, Sauer F, Kuwert T and Xu C 2006 A practical salient region feature based 3d multi-modality registration method for medical images *Proc. SPIE on Medical Imaging,* vol 6144 ed J Reinhardt and J Pluim pp 870–9

[51] Ferrant M, Macq B, Nabavi A, Jolesz F A, Kikinis R and Warfield S K 2001 Registration of 3-d intraoperative MR images of the brain using a finite-element biomechanical model *IEEE TMI* **20** 1384–97

[52] Sermesant M, Forest C, Pennec X, Delingette H and Ayache N 2003 Deformable biomechanical models: application to 4D cardiac image analysis *Med. Image Anal.* **7** 475–88

[53] Roche A 2001 Recalage d'images médicales par inférence statistique *PhD Thesis* Université de Nice, Sophia-Antipolis, France

[54] Hermosillo G 2002 Variational methods for multimodal image matching *PhD Thesis* Université de Nice, France

[55] Pluim J P W, Maintz J B A and Viergever M A 1999 Mutual-information-based registration of medical images: a survey *IEEE Trans. Med. Imaging* **22** 986–1004

[56] Pluim J P W, Maintz J B A and Viergever M A 1999 Interpolation artefacts in mutual information based image registration *Proc. SPIE 2004, Medical Imaging, 1999* vol 3661 ed K M Hanson (Bellingham, WA: SPIE) pp 56–65

[57] Fischer B and Modersitzki J 2002 Fast diffusion registration *AMS Contemporary Mathematics, Inverse Problems, Image Analysis, and Medical Imaging* **313** 117–29

[58] Fischer B and Modersitzki J 2003 Curvature based image registration *J. Math. Imag. Vis.* **18** 81–5

[59] Henn S 2006 A multigrid method for a fourth-order diffusion equation with application to image processing *SIAM J. Sci. Comput.* **27** 3

[60] Haber E and Modersitzki J 2006 A multilevel method for image registration *SIAM J. Sci. Comput.* **27** 1594–607

[61] Johnson H J and Christensen G E 2002 Consistent landmark and intensity-based image registration *IEEE Trans. Med. Imaging* **21** 450–61

[62] Rohlfing T, Maurer C R, Bluemke D A and Jacobs M A 2003 Volume-preserving nonrigid registration of MR breast images using free-form deformation with an incompressibility constraint *IEEE Trans. Med. Imaging* **22** 730–41

[63] Little J A, Hill D L G and Hawkes D J 1997 Deformations incorporating rigid structures *Comput. Vis. Image Understand.* **66** 223–32

[64] Loeckx D, Maes F, Vandermeulen D and Suetens P 2004 Nonrigid image registration using free-form deformations with a local rigidity constraint *MICCAI (LNCS vol 3216)* pp 639–46

[65] Staring M, Klein S and Pluim J P W 2006 Nonrigid registration using a rigidity constraint *Proc. SPIE 2006: Medical Imaging* vol 6144 ed J M Reihnardt and J P W Pluim (Bellingham, WA: SPIE) pp 1–10

[66] Modersitzki J 2008 FLIRT with rigidity—image registration with a local non-rigidity penalty *Int. J. Comput. Vis.* 153–63

[67] Miller M, Trouvé A and Younes L 2002 On the metrics and Euler–Lagrange equations of computational anatomy *Annu. Rev. Biomed. Eng.* **4** 375–405

[68] Beg F, Miller M, Trouvé A and Younes L 2005 Computing large deformation metric mappings via geodesic flows of diffeomorphisms *Int. J. Comput. Vis.* **61** 139–57

[69] Rueckert D, Aljabar P, Heckemann R A, Hajnal J V and Hammers A 2006 Diffeomorphic registration using B-Splines *9th Int. Conf. Medical Image Computing and Computer-Assisted Intervention (MICCAI 2006)*

[70] Fischer B and Modersitzki J 2003 Combining landmark and intensity driven registrations *Proc. Appl. Math. Mech.* **3** 32–5

[71] Haber E and Modersitzki J 2007 COFIR: coarse and fine image registration *Real-Time PDE-Constrained Optimization* ed L T Biegler, O Ghattas, M Heinkenschloss, D Keyes and B van Bloemen Waanders (Philadelphia, PA: SIAM) pp 37–49

[72] Haber E, Heldmann S and Modersitzki J 2007 Image registration constrained by local rigidity *SIIMS* 1–16 submitted

[73] Avants B B, Schoenemann P T and Gee J C 2006 Lagrangian frame diffeomorphic image registration: Morphometric comparison of human and chimpanzee cortex *Med. Image Anal.* **10** 397–412

[74] Fletcher R 1976 Conjugate gradient methods for indefinite systems *Numerical Analysis—Dundee 1975 (Lecture Notes in Mathematics* vol 506) ed G A Watson (Heidelberg: Springer) pp 73–89

[75] Gill P E, Murray W and Wright M H 1981 *Practical Optimization* (London: Academic)

[76] Dennis J E and Schnabel R B 1996 Numerical methds for unconstrained optimization and nonlinear equations *Classics in Applied Mathematics* (Philadelphia, PA: SIAM)

[77] Nocedal J and Wright S J 1999 *Numerical Optimization* (New York: Springer)

[78] Bro-Nielsen M and Gramkow C 1996 Fast fluid registration of medical images *Lecture Notes Comput. Sci.* **1131** 267–76

[79] Henn S and Witsch K 2003 Multi-modal image registration using a variational approach *SIAM J. Sci. Comput.* **25** 1429–47

[80] Henn S and Witsch K 2000 A multigrid approach for minimizing a nonlinear functional for digital image matching *Computing* **64** 339–48

[81] Henn S and Witsch K 2001 Iterative multigrid regularization techniques for image matching *SIAM J. Sci. Comput.* **23** 1077–93

[82] Schmitt O. 2001 *Die multimodale Architektonik des menschlichen Gehirns* (Habilitation, Germany: Insitute of Anatomy, Medical University of Lübeck)

[83] Wyatt P P, Noble J A, Burcher M and English R E 2002 MAP MRF joint segmentation and registration *MICCAI 2002 (Tokyo)* pp 580–7

[84] Modersitzki J and Wirtz S 2006 Combining homogenization and registration *Biomedical Image Registration: Third Int. Workshop, WBIR 2006 (LNCS)* ed J P W Pluim, Boštjan Likar and F A Gerritsen (Berlin: Springer) pp 257–63