Int J Comput Vis (2008) 76: 153–163
DOI 10.1007/s11263-007-0079-3

POSITION PAPER

FLIRT with Rigidity—Image Registration with a Local Non-rigidity Penalty

Jan Modersitzki

Received: 24 November 2006 / Accepted: 23 July 2007 / Published online: 15 September 2007

Abstract Registration is a technique nowadays commonly used in medical imaging. A drawback of most of the current registration schemes is that all tissue is being considered as non-rigid (Staring et al., Proceedings of the SPIE 2006, vol. 6144, pp. 1–10, 2006). Therefore, rigid objects in an image, such as bony structures or surgical instruments, may be transformed non-rigidly. In this paper, we integrate the concept of local rigidity to the FLexible Image Registration Toolbox (FLIRT) (Haber and Modersitzki, in SIAM J. Sci. Comput. 27(5):1594–1607, 2006; Modersitzki, Numerical Methods for Image Registration, 2004). The idea is to add a penalty for local non-rigidity to the cost function and thus to penalize non-rigid transformations of rigid objects. As our examples show, the new approach allows the maintenance of local rigidity in the desired fashion. For example, the new scheme can keep bony structures rigid during registration.

We show, how the concept of local rigidity can be integrated in the FLIRT approach and present the variational backbone, a proper discretization, and a multilevel optimization scheme. We compare the FLIRT approach to the B-spline approach. As expected from the more general setting of the FLIRT approach, our examples demonstrate that the FLIRT results are superior: much smoother, smaller deformations, visually much more pleasing.

Keywords Image processing · Image registration · Warping · Fusion · Rigidity constraints · Variational techniques · Constrained optimization

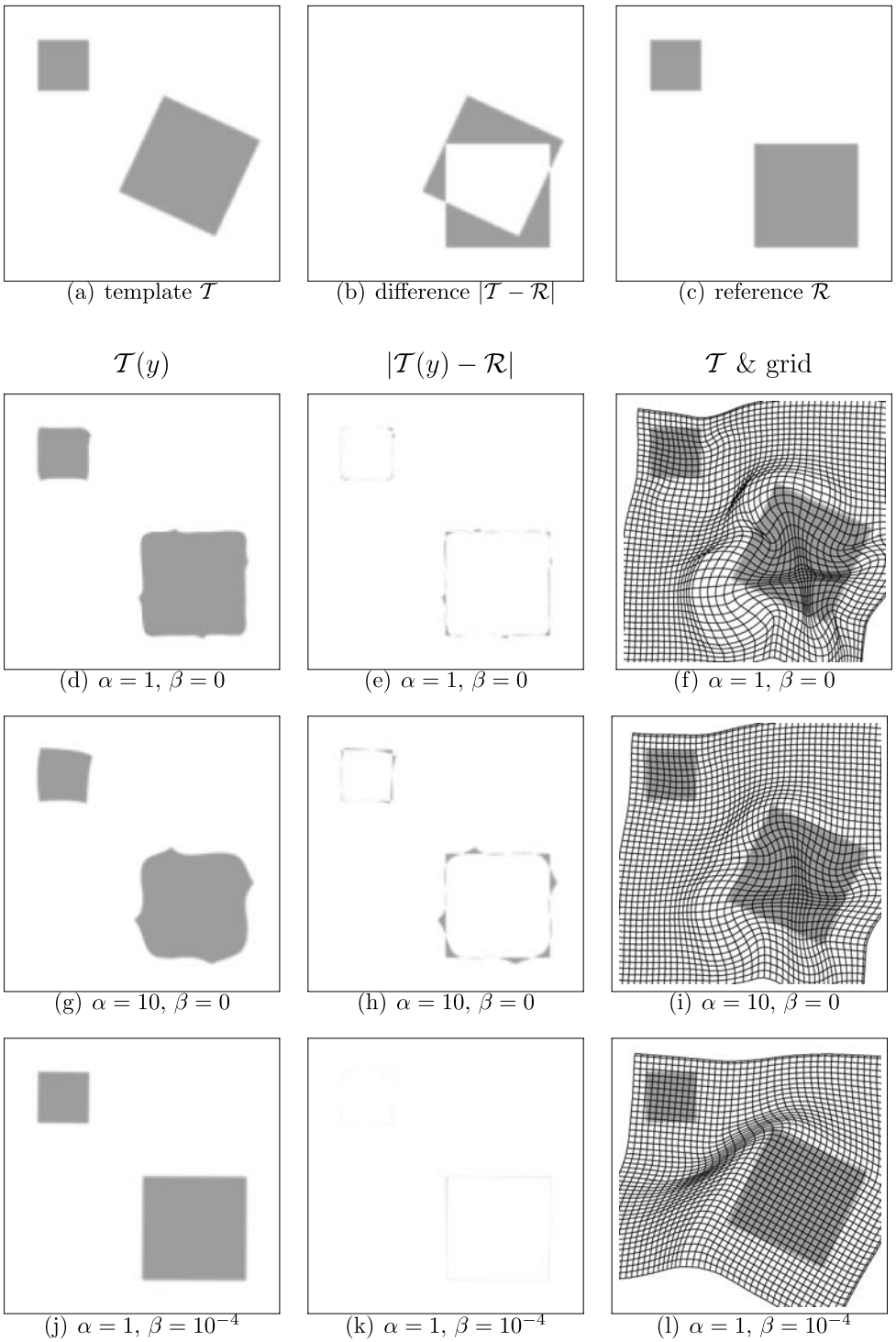
J. Modersitzki (✉)
Institute of Mathematics, University of Lübeck, Wallstraße 40, 23560 Lübeck, Germany
e-mail: modersitzki@math.uni-luebeck.de

1 Introduction

1.1 Medical Image Registration

Registration is the determination of a reasonable, geometrical transformation that aligns points in one view of an object with corresponding points in another view of the same or a similar object. There exist many instances in imaging which demand for registration, particularly in medical imaging. Examples include the treatment verification of pre- and post-intervention images, the study of temporal series of cardiac images, and the monitoring of the time evolution of an agent injection subject to patient motion. Another important application is the combination of information from multiple images, acquired using different modalities, like for example computer tomography (CT) and magnetic resonance imaging (MRI), a technique also known as fusion. The problem of fusion and registration arises whenever images acquired from different subjects, at different times, or from different scanners need to be combined for analysis or visualization. In the last two decades, computerized non-rigid image registration has played an increasingly important role in medical imaging; see, e.g., Maintz and Viergever (1998), Glasbey (1998), Peckar et al. (1999), Fitzpatrick et al. (2000), Hajnal et al. (2001), Zitová and Flusser (2003), Yoo (2004), Modersitzki (2004) and references therein.

Although registration is a generic image processing task like segmentation or enhancement, the tools to be used strongly depend on the particular application. The main difficulty in registration arises from ill-posedness (Droske and Rumpf 2004; Modersitzki 2004): for every point in a view we give one information (grayvalue) but ask for two (or even three) unknowns (the transformation). Therefore, regularization becomes a key issue. Figure 1 shows reference



(a) template $\mathcal{T}$ (b) difference $|\mathcal{T} - \mathcal{R}|$ (c) reference $\mathcal{R}$

(d) $\alpha = 1, \beta = 0$ (e) $\alpha = 1, \beta = 0$ (f) $\alpha = 1, \beta = 0$

(g) $\alpha = 10, \beta = 0$ (h) $\alpha = 10, \beta = 0$ (i) $\alpha = 10, \beta = 0$

(j) $\alpha = 1, \beta = 10^{-4}$ (k) $\alpha = 1, \beta = 10^{-4}$ (l) $\alpha = 1, \beta = 10^{-4}$

Fig. 1 Registration results for $\mathcal{R}$ (**c**) and $\mathcal{T}$ (**a**); transformed template $\mathcal{T}(y)$ (*first column*), difference $|\mathcal{T}(y) - \mathcal{R}|$ (*second column*), and visualization of y (*third column*)

and template images and visualizes three different transformations, obtained by weak regularization, strong regularization, and low regularization with additional penalty. Although the transformations are quite different, the transformed images almost look the same. Without additional, external knowledge, it is impossible to rank these transformations. More recent approaches to registration therefore aim to incorporate additional pre-knowledge, as for example the correspondence of anatomical landmarks (Johnson and Christensen 2002; Fischer and Modersitzki 2003a, 2003b), volume preservation of tissue (Rohlfing et al. 2003; Haber and Modersitzki 2004), limitation of volume changes (Haber and Modersitzki 2006b), or local rigidity (Staring et al. 2006; Modersitzki 2007).

1.2 Local Rigidity

Most of the current registration schemes (except for Little et al. 1997) treat all displayed tissue either as globally rigid or globally non-rigid (Staring et al. 2006). Therefore, rigid objects in an image, such as bony structures or surgical instruments, may be non-rigidly deformed. Other consequences are that tumor growth between follow-up images may be concealed, or that structures containing contrast material in only one of the images may be compressed by the registration scheme. In order to improve the registration results, we thus aim to integrate a concept of local rigidity as illustrated in Fig. 1. Reference and template images provide a natural segmentation in two gray squares on a white background. The new objective is to find a transformation maintaining the inner structure of the segments, i.e., being locally rigid. In a medical application, the concept of local rigidity applies to perfection to the registration of bones but can also lead to more realistic results for the generic case of images displaying soft and hard tissues.

1.3 State of the Art

There are currently two main numerical approaches to image registration. The first one is based on an expansion of the wanted transformation in terms of B-splines (Rueckert et al. 1999) and the other one is based on the more general variational framework (Modersitzki 2004) (FLexible Image Registration Toolbox (FLIRT)). Both approaches principally allow for the integration of additional pre-knowledge in terms of a penalty, like, e.g. local rigidity. For a B-spline or an even more general radial basis function approach this has been implemented in (Little et al. 1997; Loeckx et al. 2004; Staring et al. 2006), while the objective of this paper is the integration of the local rigidity concept into FLIRT.

In contrast to (Keeling and Ring 2005), where the regularizer is constructed to penalize departure from rigidity on the global average, our approach penalizes departure from rigidity on average over locally selected regions; see in particular Fig. 9 in (Keeling and Ring 2005). Here, we do not use a spatially variant regularization parameter (e.g., local elasticity parameters like Kabus et al. 2006) but explicitly penalize local non-rigidity.

1.4 FLIRT

The FLexible Image Registration Toolbox (FLIRT) (Fischer and Modersitzki 2003c; Modersitzki 2004, and in particular Haber and Modersitzki 2004, 2006a, 2006b) provides a general and efficient framework for image registration. This approach treats registration as an optimization problem where the objective is to minimize a joint functional $\mathcal{J}$ with respect to the wanted transformation y. This functional is assembled from different building blocks: a distance measure $\mathcal{D}$ quantifying image similarity, a regularizer $\mathcal{S}$ quantifying reasonability of the transformation, a penalty $\mathcal{P}$ penalizing unwanted transformations, and even additional hard constraints, if wanted. Our numerical approach uses an efficient multilevel strategy which combines different discretizations of the continuous problem. On a fixed level, a Gauss-Newton type optimization scheme is used, where a multigrid scheme serves as a solver for the linear system.

1.5 Outline

In this paper, we integrate the concept of local rigidity to the FLexible Image Registration Toolbox (FLIRT) (Fischer and Modersitzki 2003c; Modersitzki 2004). The idea is to add a penalty for local non-rigidity to the cost function and thus to penalize non-rigid transformations of rigid objects. In Sect. 2 we briefly summarize the FLIRT approach and present a variational setting for local rigidity, whereas Sect. 3 summarizes the discretization of the variational FLIRT approach and presents our discretization of local rigidity. The optimization scheme is explained in Sect. 4. In Sect. 5 we present numerical results. We demonstrate that the new approach enables the maintenance of local rigidity. Moreover, we show that the FLIRT approach is superior to the B-spline approach (as expected from the additional smoothing properties of B-splines). We finally show the performance in a medical setting: in an X-ray application we keep a middle finger of a hand rigid and in a MRI application we keep the bones in a leg rigid.

2 Mathematical Framework

We briefly summarize the variational framework of the FLIRT approach (see Haber and Modersitzki 2004, 2006a, 2006b for details). Moreover, we also present a functional setting for the local rigidity approach. The basic idea is to measure rigidity (by means of linearity, orthogonality and orientation preservation of the transformation) at every spatial position and to finally weight this measure by a moving weighting function $\mathcal{Q}(y)$.

2.1 Summarizing FLIRT

The FLIRT approach is based on the minimization of the joint functional

$$\mathcal{J}(y) := \mathcal{D}(\mathcal{T}(y), \mathcal{R}) + \alpha \mathcal{S}(y - y^{\text{kern}}) + \beta \mathcal{C}(y) \quad \text{subject to} \quad y \in \mathcal{M}. \tag{1}$$

Here, $y = (y_1, \ldots, y_d) : \mathbb{R}^d \to \mathbb{R}^d$ is the wanted transformation, $\mathcal{R}, \mathcal{T} : \Omega \to \mathbb{R}$ are the given reference and template images, and $\mathcal{T}(y)$ is the transformed image, i.e.,

$$\mathcal{T}(y)(x) := \mathcal{T}(y(x)) \quad \text{for all } x \in \Omega, \tag{2}$$

where $\Omega \subset \mathbb{R}^d$ is a given domain (typically $\Omega =]0, 1[^d$) and d is the image dimensionality (typically $d = 2, 3$).

The functional $\mathcal{D}$ measures the distance between the images and could be the sum of squares difference (SSD), mutual information (Collignon et al. 1995; Wells et al. 1996), normal gradient field (Pluim et al. 2000; Haber and Modersitzki 2005), or any other distance measure at hand. In this paper, we focus on SSD,

$$\mathcal{D}(\mathcal{T}(y), \mathcal{R}) = \frac{1}{2}\|\mathcal{T}(y) - \mathcal{R}\|^2_{L_2(\Omega)}. \tag{3}$$

The regularizer $\mathcal{S}$ is designed to yield a unique transformation y. Common choices are summarized in (Modersitzki 2004). Here, we use the most common so-called elastic potential

$$\mathcal{S}(y) = \frac{\omega}{2}\sum_j \|\nabla y^j\|^2_{L_2(\Omega)} + \frac{\lambda+\omega}{2}\|\nabla \cdot y\|^2_{L_2(\Omega)}, \tag{4}$$

with Lamé-constants λ, ω; see Modersitzki (2004) for details. The regularization parameter $\alpha > 0$ compromises between similarity and regularity. Note that the regularization in (1) is applied to the displacement $y - y^{\text{kern}}$ which allows an elegant integration of any pre-registration. The constraints $\mathcal{C}$ and $\mathcal{M}$ are used to supply additional application conform pre-knowledge. Particular choices for local rigidity are discussed in the next subsection.

2.2 Local Rigidity

The goal of this subsection is to formalize and facilitate local rigidity. The naive approach would be based on a segmentation and may lead to localization errors in a multi-level framework. We therefore prefer a weighted approach, where the penalty is computed everywhere and then weighted by a differentiable function $\mathcal{Q} : \Omega \to \mathbb{R}$; see Fig. 4j and Fig. 5j for examples. The idea is to segment structures that are supposed to transform rigidly and to pick Q as a smoothed characteristic function for the a priori given segments. In the examples shown, a manual segmentation has been performed and Q is based on a linear interpolation scheme. Note that the rigid transformations on different segments can be different as it is shown in the knee example. A second advantage of the weighted approach is that it easily allows for a correct location of the penalty in a multilevel framework. Since we like to maintain the rigidity of a displayed structure and not of a fixed subset of the domain Ω, the weight has to be transformed in the same way as the structure. As a consequence, the appropriate weight is $\mathcal{Q}(y)$, where $\mathcal{Q}(y)(x) = \mathcal{Q}(y(x))$.

We start with the Jacobian of y,

$$\nabla y = \begin{pmatrix} \partial_1 y_1 & \ldots & \partial_d y_1 \\ \vdots & \ldots & \vdots \\ \partial_1 y_d & \ldots & \partial_d y_d \end{pmatrix}, \tag{5}$$

where natural boundary conditions associated with ∇y are used, i.e., that no additional artificial boundary condition constraints are imposed. A formal definition of rigidity reads as follows, see also Gurtin (1981), Staring et al. (2006).

Definition 1 *A transformation $y \in C^2(\Omega, \mathbb{R}^d)$ is rigid, if it is linear ($\partial_{i_1,i_2} y_j = 0$ for all $i_1, i_2, j = 1, \ldots, d$), orthogonal ($\nabla y^\top \nabla y = I_d$), and orientation preserving ($\det \nabla y = 1$) for all $x \in \Omega$.*

Lemma 1 *A transformation $y \in C^2(\Omega, \mathbb{R}^d)$ is rigid, if and only if $y(x) = Qx + b$ where $Q \in \mathbb{R}^{d,d}$ is orthogonal and orientation preserving, and $b \in \mathbb{R}^d$.*

Proof From $\partial_{i_1,i_2} y_j = 0$ it follows that the transformation is linear, i.e. $y(x) = Ax + b$ and $\nabla y(x) = A$, from $\nabla y^\top \nabla y = I_d$ it follows that A is orthogonal and from $\nabla y^\top \nabla y = I_d$ orientation preservation follows. The opposite direction is obvious. □

With this definition, we have $\mathcal{M} = C^2(\Omega, \mathbb{R}^d)$ and the penalty $\mathcal{C}$ in (1) is given by

$$\begin{aligned}\mathcal{C}(y) = &\sum_{i_1,i_2=1,i_2\le i_1}^{d} \frac{1}{2}\|\partial_{i_1,i_2} y_j\|^2_{\mathcal{Q}(y)} \\ &+ \sum_{i_1,i_2=1,i_2\le i_1}^{d} \frac{1}{2}\|(\nabla y^\top \nabla y - I_d)_{i_1,i_2}\|^2_{\mathcal{Q}(y)} \\ &+ \frac{1}{2}\|\det \nabla y - 1\|^2_{\mathcal{Q}(y)},\end{aligned} \tag{6}$$

with the weighted L_2-norm

$$\|f\|^2_{\mathcal{Q}(y)} = \int_\Omega \mathcal{Q}(y(x))^2\, f(x)^2\, dx. \tag{7}$$

The first term in (6) measures departure from linearity, the second departure from orthogonality, and the third departure from orientation preservation; the weighted norm puts a focus on the moving structures to be kept rigid.

Particularly for $d = 2$, and $\|\cdot\| = \|\cdot\|_{\mathcal{Q}(y)}$ we have

$$\begin{aligned}2\mathcal{C}(y) = &\|\partial_{1,1} y_1\|^2 + \|\partial_{1,2} y_1\|^2 + \|\partial_{2,2} y_1\|^2 \\ &+ \|\partial_{1,1} y_2\|^2 + \|\partial_{1,2} y_2\|^2 + \|\partial_{2,2} y_2\|^2 \\ &+ \|(\partial_1 y_1)^2 + (\partial_2 y_1)^2 - 1\|^2 \\ &+ \|(\partial_1 y_2)^2 + (\partial_2 y_2)^2 - 1\|^2\end{aligned}$$

$$+\|(\partial_1 y_1)(\partial_1 y_2)+(\partial_2 y_1)(\partial_2 y_2)\|^2$$
$$+\|(\partial_1 y_1)(\partial_2 y_2)-(\partial_2 y_1)(\partial_1 y_2)-1\|^2, \quad (8)$$

and for $d=3$ the formula is along the same lines but lengthy.

3 Discretization

Our numerical approach focusses on a discretize then optimize strategy, see Haber and Modersitzki (2004), Haber and Modersitzki (2006a) for details. The discretization is based on pixel/voxel (see Fig. 2). An important issue is to guarantee an h-elliptic discretization for the regularizer. We therefore use staggered grids for the discretization of the vector field y. We describe the discretization for $d=2$, the extension to $d=3$ is straightforward; more details are given in Haber and Modersitzki (2004). In Sect. 3.1 we briefly summarize the FLIRT discretization (Haber and Modersitzki 2004, 2006a) and in Sects. 3.2 and 3.3 we explain the discretization of the rigidity penalty in detail.

3.1 Discretizing the Images, the Distance, and the Regularizer

We briefly summarize the discretization suggested in Haber and Modersitzki (2004) for dimension $d=2$. For the discretization of y we assume that $\Omega=]0,\omega_1[\times]0,\omega_2[$ and $m=(m_1,m_2)$ are given. For any dimension j, we set $h_j=\omega_j/m_j$ and end up with the following grids (see also Fig. 2):

$$x^{h,\bullet}=(i_1h_1,i_2h_2)_{i_j=1/2:m_j-1/2}, \quad \text{cell-centered},$$
$$x^{h,\blacktriangleright}=(i_1h_1,i_2h_2)_{i_1=0:m_1,i_2=1/2:m_2-1/2}, \quad \text{face staggered},$$
$$x^{h,\blacktriangle}=(i_1h_1,i_2h_2)_{i_1=1/2:m_1-1/2,i_2=0:m_2}, \quad \text{face staggered}.$$

We discretize $y=(y_1(x_1,x_2),y_2(x_1,x_2))$ by values on the staggered grids, i.e. $y_1^h=y^1(x^{h,\blacktriangleright})$ and $y_2^h=y^1(x^{h,\blacktriangle})$, where for convenience, we assemble the unknowns in lexicographical order, i.e. $y_1^h\in\mathbb{R}^{n_1}$ and $y_2^h\in\mathbb{R}^{n_2}$ with $n_1=(m_1+1)m_2$, $n_2=m_1(m_2+1)$, respectively.

The ultimate unknown is the vector $y^h=((y_1^h)^\top,(y_2^h)^\top)^\top\in\mathbb{R}^n$, with $n=n_1+n_2$. Explicit formulae for the discretization of $\mathcal{T}$, $\mathcal{R}$, $\mathcal{D}$ and $\mathcal{S}$ are given in Haber and Modersitzki (2004) and are summarized as follows:

$$R=\mathcal{R}(x^{h,\bullet})\in\mathbb{R}^n,$$
$$T(y^h)=\mathcal{T}(Py^h)\in\mathbb{R}^n,$$
$$D(T(y^h),R)=\frac{1}{2}\|T(y^h)-R\|^2,$$
$$S(y^h)=\frac{1}{2}\|By^h\|^2,$$

where

$$P=\begin{pmatrix}P_1 & 0\\ 0 & P_2\end{pmatrix},$$
$$\|By^h\|^2=\frac{\lambda+\omega}{2}\|\nabla^h\cdot y^h\|^2+\frac{\omega}{2}\sum_{j=1}^{d}\|\nabla_j^h y_j^h\|^2,$$
$$\nabla^h\cdot=(\partial_1^{h,1},\partial_2^{h,2}),\qquad \nabla_j^h=((\partial_1^{h,j})^\top,(\partial_2^{h,j})^\top)^\top,$$

and the matrices P_1, P_2, and $\partial_k^{h,j}$ are specified in the next section.

3.2 Discretizing Derivatives

For the regularizer and the rigidity constraints, we need to discretize first and second order derivatives of y. The goal and the reason for staggered grids is to use short differences for the approximations of $\partial_j y_j$, which are then located at the cell-centered grid. Unfortunately, all other derivatives involve averaging operators that keep a formal description lengthy. Our description is based on the band matrices and Kronecker-products $\otimes$ (Brewer 1978). Let band$(a_{-k},\ldots,a_k;k_1,k_2)$ denotes a k_1-by-k_2 band matrix with diagonal bands $a_{-k},\ldots,a_k$, where $\underline{a_0}$ is on the main diagonal. For example,

$$\text{band}(0,\underline{-1},1;3,4)=\begin{pmatrix}\underline{-1} & 1 & 0 & 0\\ 0 & \underline{-1} & 1 & 0\\ 0 & 0 & \underline{-1} & 1\end{pmatrix}.$$

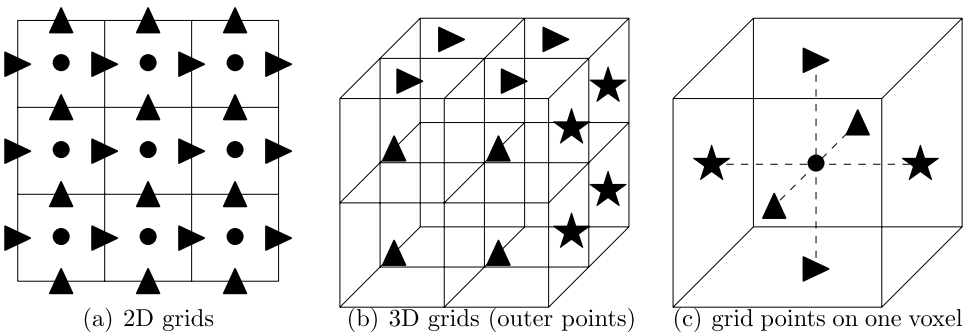


Fig. 2 Staggered grid discretization of y

(a) 2D grids (b) 3D grids (outer points) (c) grid points on one voxel

With these matrices, we set

$$
\begin{aligned}
&P_j = \text{band}(0, \underline{1}, 1; m_j, m_j+1)/2,\\
&D^{\text{short},j} = \text{band}(0, \underline{-1}, 1; m_j, m_j+1)/h_j,\\
&D^{\text{long},j} =_{\text{BC}} \text{band}(-1, \underline{0}, 1; m_j, m_j+1)/(2h_j),\\
&D^{\text{short},j,j} = \text{band}(1, \underline{-2}, 1; m_j, m_j)/(h_j^2),\\
&D^{\text{long},j,j} =_{\text{BC}} \text{band}(0, 1, \underline{-1}, -1, 1; m_j, m_j+1)/(2h_j^2),
\end{aligned}
$$

where "$=_{\text{BC}}$" means "up to boundary conditions". Particularly, the boundary conditions are chosen such that discretizations of constants (for first order derivatives) and affine linear (second order derivatives) belong to the kernel.

The first and second order derivatives (linear parts in (6)) are approximated by

$$\partial_1 y_1(x^{h,\bullet}) \approx \partial_1^{h,1} y_1^h, \quad \partial_1^{h,1} = I_{m_2} \otimes D^{\text{short},1}, \tag{9a}$$

$$\partial_2 y_1(x^{h,\bullet}) \approx \partial_2^{h,1} y_1^h, \quad \partial_2^{h,1} = D^{\text{long},2} \otimes P_1, \tag{9b}$$

$$\partial_1 y_2(x^{h,\bullet}) \approx \partial_1^{h,2} y_2^h, \quad \partial_1^{h,2} = P_2 \otimes D^{\text{long},1}, \tag{9c}$$

$$\partial_2 y_2(x^{h,\bullet}) \approx \partial_2^{h,2} y_2^h, \quad \partial_2^{h,2} = D^{\text{short},2} \otimes I_{m_1}, \tag{9d}$$

$$\partial_{1,1} y_1(x^{h,\bullet}) \approx C_1(y^h) = \partial_{1,1}^{h,1} y_1^h, \quad \partial_{1,1}^{h,1} = I_{m_2} \otimes D^{\text{long},1,1}, \tag{10a}$$

$$\partial_{1,2} y_1(x^{h,\bullet}) \approx C_2(y^h) = \partial_{1,2}^{h,1} y_1^h, \quad \partial_{1,2}^{h,1} = D^{\text{long},2} \otimes D^{\text{short},1}, \tag{10b}$$

$$\partial_{2,2} y_1(x^{h,\bullet}) \approx C_3(y^h) = \partial_{2,2}^{h,1} y_1^h, \quad \partial_{2,2}^{h,1} = D^{\text{short},2,2} \otimes P_1, \tag{10c}$$

$$\partial_{1,1} y_2(x^{h,\bullet}) \approx C_4(y^h) = \partial_{1,1}^{h,2} y_2^h, \quad \partial_{1,1}^{h,2} = P_2 \otimes D^{\text{short},1,1}, \tag{10d}$$

$$\partial_{1,2} y_2(x^{h,\bullet}) \approx C_5(y^h) = \partial_{1,2}^{h,2} y_2^h, \quad \partial_{1,2}^{h,2} = D^{\text{short},2} \otimes D^{\text{long},1}, \tag{10e}$$

$$\partial_{2,2} y_2(x^{h,\bullet}) \approx C_6(y^h) = \partial_{2,2}^{h,2} y_2^h, \quad \partial_{2,2}^{h,2} = D^{\text{long},1,1} \otimes I_{m_1}, \tag{10f}$$

and the non-linear part $(\nabla y^\top \nabla y - I_d)$ and $(\det \nabla y - 1)$ in (6) are discretized by

$$C_7(y^h) = (\partial_1^{h,1} y_1^h) \odot (\partial_1^{h,1} y_1^h) + (\partial_2^{h,1} y_1^h) \odot (\partial_2^{h,1} y_1^h) - e, \tag{10g}$$

$$C_8(y^h) = (\partial_1^{h,1} y_1^h) \odot (\partial_1^{h,2} y_2^h) + (\partial_2^{h,1} y_1^h) \odot (\partial_2^{h,2} y_2^h), \tag{10h}$$

$$C_9(y^h) = (\partial_1^{h,2} y_2^h) \odot (\partial_1^{h,2} y_2^h) + (\partial_2^{h,2} y_2^h) \odot (\partial_2^{h,2} y_2^h) - e, \tag{10i}$$

$$C_{10}(y^h) = (\partial_1^{h,1} y_1^h) \odot (\partial_2^{h,2} y_2^h) - (\partial_2^{h,1} y_1^h) \odot (\partial_1^{h,2} y_2^h) - e, \tag{10j}$$

where $e = (1, \ldots, 1)^\top$ and $\odot$ denotes the Hadamard- or pointwise product ($(x \odot y)_i = x_i y_i$).

3.3 Discretizing Local Rigidity

Using the discretizations of the previous subsection, we discretize the penalty $\mathcal{C}$ by an l_2-norm of an appropriate chosen residual r:

$$C(y^h) = \frac{1}{2} r(y^h)^\top r(y^h), \quad r = (r_1^\top, \ldots, r_{10}^\top)^\top,$$
$$\text{and } r_j(y^h) = \mathcal{Q}(P y^h) \odot C_j(y^h). \tag{11}$$

For the derivative of C, we thus get

$$dC(y^h) = r(y^h)^\top dr(r^h), \tag{12}$$

where the derivative dr_j is given by

$$dr_j(y^h) = \text{diag}(Q(Py^h)) dC_j(y^h) + \text{diag}(C_j(y^h)) d\mathcal{Q}(Py^h) P. \tag{13}$$

Note that for $j = 1, \ldots, 6$, C_j is linear and dC_j follows directly from (9). The parts $C_7, \ldots, C_{10}$ are quadratic and their derivatives are given by

$$dC_7(y^h) = (2\text{diag}(\partial_1^{h,1} y_1^h)\partial_1^{h,1} + 2\text{diag}(\partial_2^{h,1} y_1^h)\partial_2^{h,1}, 0) \tag{14a}$$

$$dC_8(y^h) = (\text{diag}(\partial_1^{h,2} y_2^h)\partial_1^{h,1} + \text{diag}(\partial_2^{h,2} y_2^h)\partial_2^{h,1}, \text{diag}(\partial_1^{h,1} y_1^h)\partial_1^{h,2} + \text{diag}(\partial_2^{h,1} y_1^h)\partial_2^{h,2}), \tag{14b}$$

$$dC_9(y^h) = (0, 2\text{diag}(\partial_1^{h,2} y_2^h)\partial_1^{h,2} + 2\text{diag}(\partial_2^{h,2} y_2^h)\partial_2^{h,2}), \tag{14c}$$

$$dC_{10}(y^h) = (\text{diag}(\partial_2^{h,2} y_2^h)\partial_1^{h,1} - \text{diag}(\partial_1^{h,2} y_2^h)\partial_2^{h,1}, \text{diag}(\partial_1^{h,1} y_1^h)\partial_2^{h,2} - \text{diag}(\partial_2^{h,1} y_1^h)\partial_1^{h,2}). \tag{14d}$$

3.4 Image Registration Discretized

To summarize, our discrete version of (1) reads:

$$\text{minimize } J^h(y^h) = \frac{1}{2}\|T(y^h) - R\|^2 + \frac{\alpha}{2}\|B(y^h - y^{\text{kern},h})\|^2 + \frac{\beta}{2}\|r(y^h)\|^2, \tag{15}$$

with analytic gradient

$$dJ^h(y^h) = (T(y^h) - R)^\top dT(y^h) + \alpha B^\top B(y^h - y^{\text{kern},h}) + \beta r(y^h)^\top dr(y^h). \tag{16}$$

Algorithm 1 (Multilevel Image Registration)

1. for level $= 0 : L$ do
2. $h = h^{(\text{level})}$;
3. if (level $= 0$), % coarsest level
4. perform rigid pre-registration and obtain $y^{(0)} = y^{\text{rigid}}$;
5. else
6. set $y^{(0)} = I_h^{2h} y^{2h,*}$;
7. endIf
8. compute minimizer $y^{h,*}$ of J^h (15) using $y^{(0)}$ as starting guess
9. endFor
10. return $y^{h,*}$.

Algorithm 2 (Minimizing $J = J^h$ (15))

1. for $k = 1, \ldots$ do
2. compute $T(y)$, $dT(y)$, $D = \|T(y) - R\|^2/2$,
   $dD = (T(y) - R)^\top dT$,
   $S = \|B(y - y^{\text{kern}})\|^2/2$, $dS = B^\top B(y - y^{\text{kern}})$,
   $C = \|r(y)\|^2/2$, $dC = r(y)^\top dr(y)$,
   compute $J = D + \alpha S + \beta C$,
   compute $dJ = dD + \alpha dS + \beta dC$;
3. check stopping criteria (see, e.g., Gill et al. 1981)
4. solve $H\delta_y = -dJ^\top$, where
   $H = \text{diag}(dT^\top dT) + \alpha B^\top B + \beta dr^\top dr$,
5. use Armijo line search to compute a "minimizer" t^* of
   $\phi(t) = J(y + t\delta_y)$,
6. update $y \leftarrow y + t^*\delta_y$;
7. endFor
8. return y.

4 Optimization

For the numerical optimization of (1) we use the techniques summarized in Haber and Modersitzki (2006a). The main ingredients are a multilevel strategy and a Gauss-Newton type optimizer for a fixed level. In the Gauss-Newton approach the Hessian is approximated by a sum of the approximations to d^2D and d^2C and the exact Hessian d^2S of the regularizer.

4.1 Multilevel

In order to compute a solution for the registration problem (1), we use a hierarchy of different discretizations $h^{(0)}, \ldots, h^{(L)}$, where $h^{(\ell+1)} = 2^{-\ell}h^{(0)}$, see Algorithm 1. Starting with the coarsest grid $\ell = 0$, we initially perform a rigid pre-registration yielding y^{kern} and compute a numerical minimizer $y^{h,*}$ of J^h, where $h = h^{(0)}$ and $y = y^{\text{kern},h}$ serves as a starting guess. If we are not yet on the finest level ($\ell < L$), we switch to a higher level ($\ell \leftarrow \ell + 1$). Starting with $y = I_h^{2h} y^{2h,*}$, where I_h^{2h} is a standard multigrid interpolation scheme (Trottenberg et al. 2001), we compute a numerical minimizer $y^{h,*}$ of J^h, where $h = h^{(\ell)}$. These steps are repeated until we reach the finest resolution $h = h^{(L)}$.

4.2 Numerical Optimization of J^h

It remains to describe our numerical scheme for the minimizing of $J = J^h$. We use a Gauss-Newton type approach where the Hessian is approximated as follows

$$d^2J(y) \approx \text{diag}(dT(y)^\top dT(y)) + \alpha B^\top B + \beta dr(y)^\top dr(y), \tag{17}$$

see Algorithm 2. Note that for the regularizer S the second order derivatives $B^\top B$ is used, while for the distance D and the penalty C only first order derivatives are used. Though the dimensions of H can be dramatically large (particularly for $d = 3$), it has only a very small number of non-zero entries per row. As a consequence, multiplication by H is an $\mathcal{O}(n)$ operation and efficient multigrid techniques might be adapted. However, in our current 2D MATLAB (MathWorks 1992) implementation, we simply use the backslash operator.

5 Results

We tested our implementation on a variety of examples. Due to page limitations, we can only present a few intuitive and representative examples.

5.1 One Square

Our first example is a repetition of the experiment performed in (Staring et al. 2006), from which we also took the results for the B-spline approaches, see Fig. 3. From these results, we see the effect of local rigidity constraints placed at non-zero locations in the moving template itself ($\mathcal{Q} = \mathcal{T}$). The B-spline implementation uses a backward interpolation scheme, while the FLIRT implementation uses a forward scheme (in fact, for the FLIRT registration we interchanged reference and template in order to make the grids comparable).

As expected, both penalized approaches keep the square rigid. However, the B-spline approach introduces large deformations in some parts of the image. This is partly due to the inappropriate vanishing Dirichlet boundary conditions (BC) but may also be related to the locations of the control points.

FLIRT takes advantage of a regularization kernel which handles global transformations. Moreover, FLIRT involves vanishing natural second order Neumann boundary conditions. As a consequence, the FLIRT result—with and without penalty—is a global rigid transformation leading to a

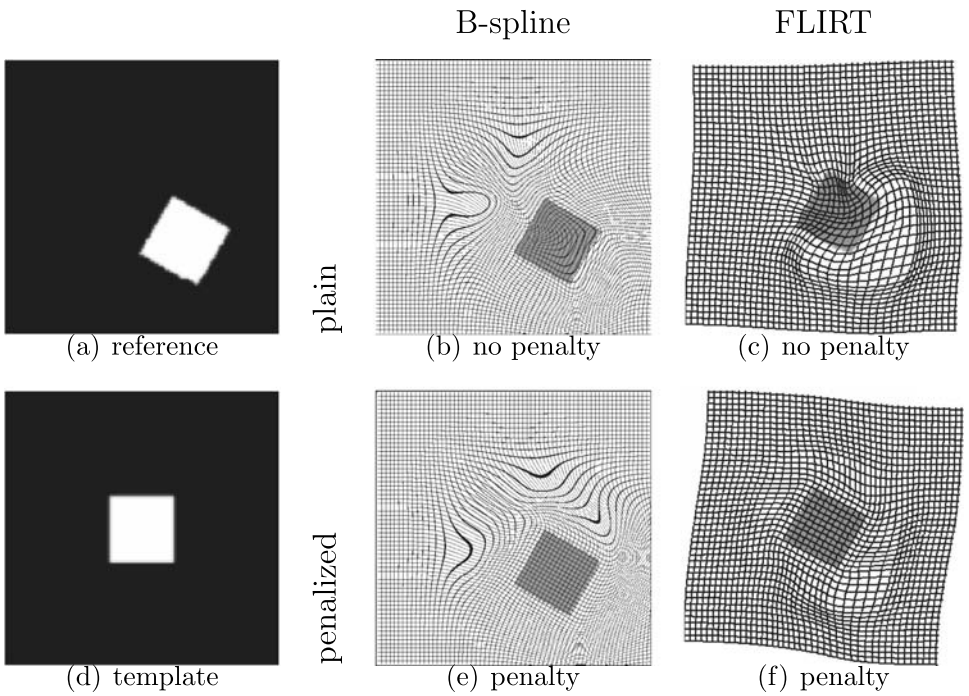


Fig. 3 Example from Staring et al. (2006): reference and template (*first column*), B-spline results taken from Staring et al. (2006) without and with penalty (*middle column*), and FLIRT results without and with penalty (*right column*); all four transformations lead to a perfect match

perfect copy of the reference. This is probably also the most intuitive solution to this toy problem. The results are not shown here. In order to show the effects of the additional penalty term, we artificially removed the kernel information and these results are shown in Fig. 3.

All results (B-spline/Flirt, without/with penalty) yield a perfect match, i.e. $D(y) \approx 0$. However, we find the FLIRT transformations superior since they are visually much smoother, more local, and introduce less distortions.

5.2 Two Squares

Our next example shows registration results for images with two rigid objects which has already been introduce in the introduction; see Fig. 1. In this example $Q = T$.

Our first experiment is done with small regularization and without penalty ($\alpha = 1$, $\beta = 0$). As a result, the transformation is heavily non-linear deformed and does not look appealing, see Fig. 1f. Therefore, the second experiment is done with large regularization but again without penalty ($\alpha = 10$, $\beta = 0$). As expected, the resulting transformation is much smoother (Fig. 1i) and looks more pleasing. However, especially the big square is transformed non-linearly and in addition, the transformed image is not even close to the reference. The third example use the local rigidity penalty ($\alpha = 1$, $\beta = 10^{-4}$). As a result, we get a very pleasing transformation (Fig. 1l) with a perfect match (Fig. 1j).

5.3 X-ray of Hands

A more realistic but still intuitive example is presented in Fig. 4 (see also Amit 1994; Modersitzki 2004). Note that the template image shows a global rotation of approximately 25 degrees which outrules the B-spline approach with Dirichlet BC. In this example we make the middle finger of the hand to be rigid (see Fig. 4j), i.e. Q is a smoothed version of the bone in the middle finger. Figure 4 shows FLIRT results without ($\beta = 0$) as well as with penalty ($\beta = 0.01$). For both variants, we picked $\alpha = 500$. As it apparent, the penalized approach does keep the finger rigid while the unconstrained does not; see particularly the maps of $\det(\nabla y)$ (see Fig. 4m, n).

5.4 Two Knees

This example shows the registration of computer tomography images from a human knee in a bent and straight position (image courtesy of T. Netsch, Philips Research Hamburg, Germany). Here, we pick two 2D subsections of the 3D reference and template images, respectively and perform the registration without ($\alpha = 200$, $\beta = 0$) and with penalty ($\alpha = 100$, $\beta = 10^{-8}$; the small value of β is related to the height weight in this example: $\max\{Q(x)\} = 255$).

For the approach without penalty, a large value $\alpha = 200$ is needed to keep the transformation one-to-one, particularly in the neighborhood of the joint. As it is apparent from Fig. 5b, even with this strong regularization the femur is intolerably deformed. For the penalized approach, we find a much smoother overall transformation (even with a smaller regularization parameter $\alpha = 100$), that leads to a realistically deformed template (Fig. 5c). The maps of $\det(\nabla y)$ (Fig. 5m, n) clearly demonstrate the benefits of local rigidity for this application.

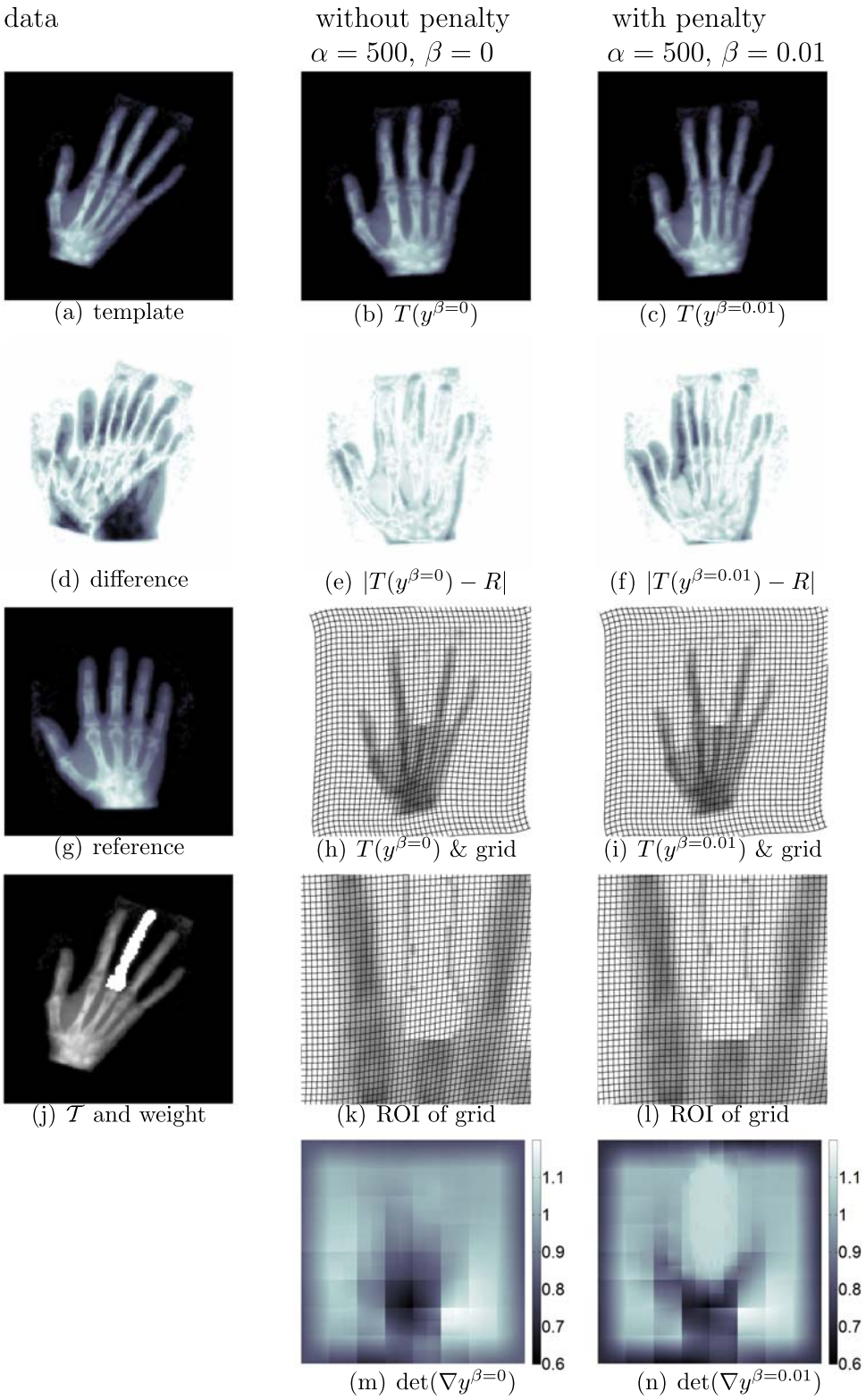


Fig. 4 Registration results for $\mathcal{R}$, (**g**) and $\mathcal{T}$, (**a**): data (*first column*), results for $\alpha = 500$ and $\beta = 0$ (*second column*) and $\alpha = 500$ and $\beta = 0.01$ (*third column*)

6 Conclusions

Registration is a technique nowadays commonly used in medical imaging. A drawback of most of the current registration schemes is that all tissue is being considered as non-rigid (Staring et al. 2006). Therefore, rigid objects in an image, such as bony structures or surgical instruments, may be transformed non-rigidly. In this paper, we integrate

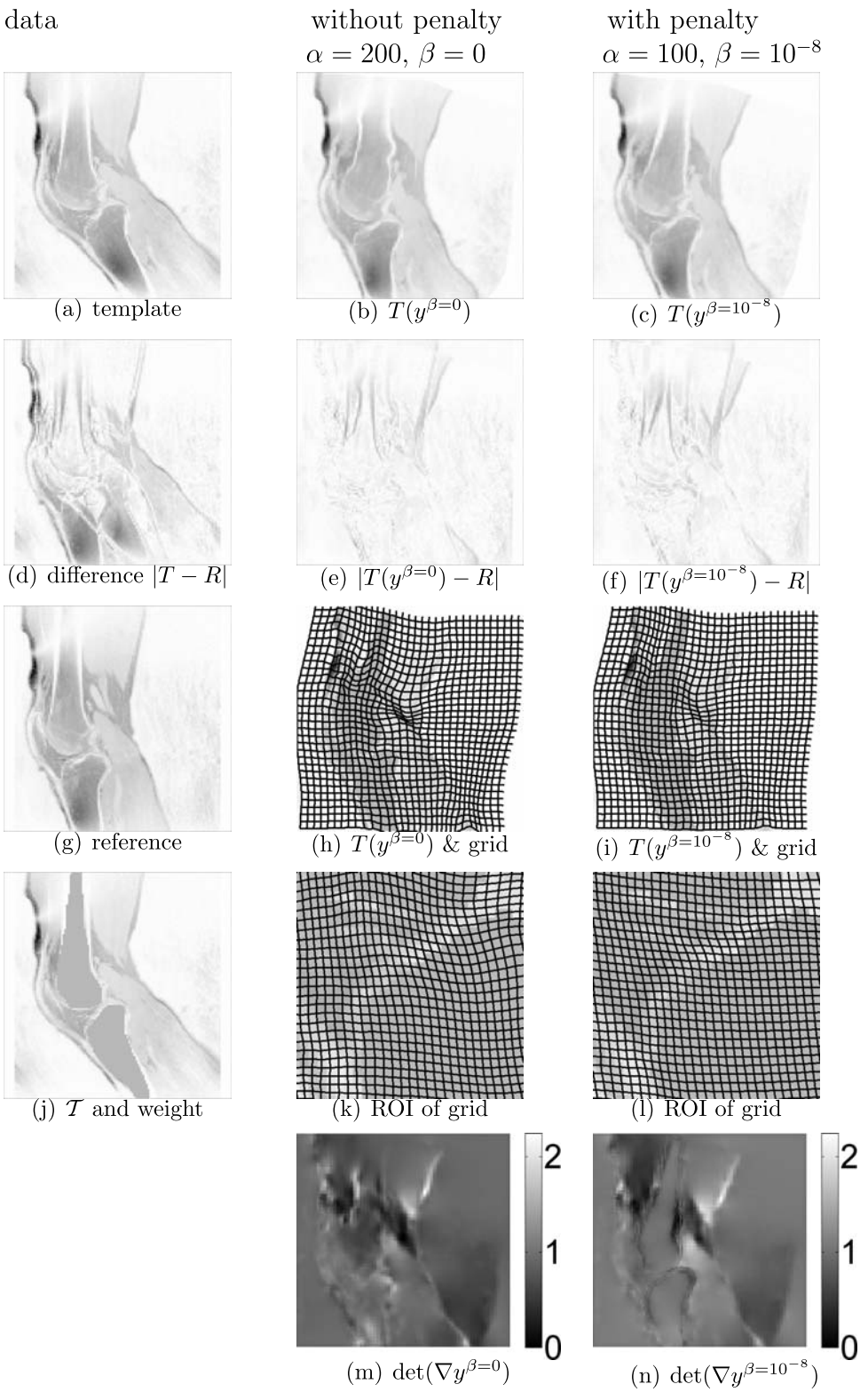


Fig. 5 Registration results for $\mathcal{R}$, (**g**) and $\mathcal{T}$, (**a**): data (*first column*), results for $\alpha = 200$ and $\beta = 0$ (*second column*) and $\alpha = 100$ and $\beta = 10^{-8}$ (*third column*)

the concept of local rigidity to the FLexible Image Registration Toolbox (FLIRT) (Modersitzki 2004; Haber and Modersitzki 2006a). The idea is to add a penalty for local non-rigidity to the cost function and thus to penalize non-rigid transformations of rigid objects. As our examples show, the new approach allows the maintenance of local rigidity in the desired fashion. For example, the new scheme can keep bony structures rigid during registration. We show, how the

concept of local rigidity can be integrated in the FLIRT approach and present the variational backbone, a proper discretization, and a multilevel optimization scheme. We compare the FLIRT approach to the B-spline approach. As expected from the more general setting of the FLIRT approach, our examples demonstrate that the FLIRT results are superior: much smoother, smaller deformations, visually much more pleasing.

Next steps are the automatization of the parameter choices for regularization and penalty and of course a 3D implementation. The FLIRT approach principally allows an integration hard constraints (like, e.g., landmarks Fischer and Modersitzki 2003b or volume preservation Haber and Modersitzki 2004). However, the integration of local rigidity as a hard constrained requires further analysis.

References

Amit, Y. (1994). A nonlinear variational problem for image matching. *SIAM Journal on Scientific Computing*, *15*(1), 207–224.

Brewer, J. W. (1978). Kronecker products and matrix calculus in system theory. *IEEE Transactions on Circuits and Systems*, *25*, 772–780.

Collignon, A., Vandermeulen, A., Suetens, P., & Marchal, G. (1995). 3d multi-modality medical image registration based on information theory. In *Computational Imaging and Vision* (Vol. 3, pp. 263–274).

Droske, M., & Rumpf, M. (2004). A variational approach to non-rigid morphological registration. *SIAM Journal on Applied Mathematics*, *64*(2), 668–687.

Fischer, B., & Modersitzki, J. (2003a). Combination of automatic non-rigid and landmark based registration: the best of both worlds. In M. Sonka & J. M. Fitzpatrick (Eds.), *Proceedings of the SPIE: Vol. 5032. Medical imaging 2003: Image processing* (pp. 1037–1048).

Fischer, B., & Modersitzki, J. (2003b). Combining landmark and intensity driven registrations. *PAMM*, *3*, 32–35.

Fischer, B., & Modersitzki, J. (2003c). FLIRT: a flexible image registration toolbox. In J. C. Gee, J. B. A. Maintz, & M. W. Vannier (Eds.), *Lecture notes in computer science: Vol. 2717. 2nd international workshop on biomedical image registration 2003* (pp. 261–270). Berlin: Springer.

Fitzpatrick, J. M., Hill, D. L. G., & Maurer Jr., C. R. (2000). Image registration. In M. Sonka, & J. M. Fitzpatrick (Eds.), *Handbook of medical imaging: Vol. 2. Medical image processing and analysis* (pp. 447–513). Bellingham: SPIE.

Gill, P. E., Murray, W., & Wright, M. H. (1981). *Practical optimization*. London: Academic Press.

Glasbey, C. (1998). A review of image warping methods. *Journal of Applied Statistics*, *25*, 155–171.

Gurtin, M. E. (1981). *An introduction to continuum mechanics*. Orlando: Academic Press.

Haber, E., & Modersitzki, J. (2004). Numerical methods for volume preserving image registration. *Inverse Problems*, *20*(5), 1621–1638.

Haber, E., & Modersitzki, J. (2005). Beyond mutual information: A simple and robust alternative. In H. P. Meinzer, H. Handels, A. Horsch, & T. Tolxdorff (Eds.), *Bildverarbeitung für die Medizin 2005* (pp. 350–354). Berlin: Springer.

Haber, E., & Modersitzki, J. (2006a). A multilevel method for image registration. *SIAM Journal on Scientific Computing*, *27*(5), 1594–1607.

Haber, E., & Modersitzki, J. (2006b). Image registration with a guaranteed displacement regularity. *International Journal of Computer Vision, 1* DOI: 10.1007/s11263-006-8984-4

Hajnal, J., Hawkes, D., & Hill, D. (2001). *Medical image registration*. Boca Raton: CRC Press.

Johnson, H. J., & Christensen, G. E. (2002). Consistent landmark and intensity-based image registration. *IEEE Transactions on Medical Imaging*, *21*(5), 450–461.

Kabus, S., Franz, A., & Fischer, B. (2006). Variational image registration with local properties. In F. A. Gerritsen, J. P. W. Pluim, & B. Likar (Eds.), *Lecture notes in computer science. Third international workshop, WBIR 2006: Biomedical image registration* (pp. 92–100). Berlin: Springer.

Keeling, S. L., & Ring, W. (2005). Medical image registration and interpolation by optical flow with maximal rigidity. *Journal of Mathematical Imaging and Vision*, *23*(1), 47–65.

Little, J. A., Hill, D. L. G., & Hawkes, D. J. (1997). Deformations incorporating rigid structures. *Computer Vision and Image Understanding*, *66*(2), 223–232.

Loeckx, D., Maes, F., Vandermeulen, D., & Suetens, P. (2004). Nonrigid image registration using free-form deformations with a local rigidity constraint. In *Lecture notes in computer science: Vol. 3216. MICCAI* (pp. 639–646).

Maintz, J. B. A., & Viergever, M. A. (1998). A survey of medical image registration. *Medical Image Analysis*, *2*(1), 1–36.

MathWorks (1992). *Matlab user's guide*. Natick.

Modersitzki, J. (2004). *Numerical methods for image registration*. Oxford: Oxford University Press.

Modersitzki, J. (2007). Image registration with local rigidity constraints. In H. Handels, H. P. Meinzer, A. Horsch, T. M. Deserno, & T. Tolxdoff (Eds.), *Informatik Aktuell: Bildverarbeitung für die Medizin* (pp. 444–448). Berlin: Springer.

Peckar, W., Schnörr, C., Rohr, K., & Stiehl, H. S. (1999). Parameter-free elastic deformation approach for 2d and 3d registration using prescribed displacements. *Journal Mathematical Imaging and Vision*, *10*(2), 143–162.

Pluim, J. P. W., Maintz, J. B. A., & Viergever, M. A. (2000). Image registration by maximization of combined mutual information and gradient information. *IEEE TMI*, *19*(8), 809–814.

Rohlfing, T., Maurer Jr., C. R., Bluemke, D. A., & Jacobs, M. A. (2003). Volume-preserving nonrigid registration of MR breast images using free-form deformation with an incompressibility constraint. *IEEE Transactions on Medical Imaging*, *22*(6), 730–741.

Rueckert, D., Sonoda, L., Hayes, C., Hill, D., Leach, M., & Hawkes, D. (1999). Non-rigid registration using free-form deformations. *IEEE Transactions on Medical Imaging*, *18*(1), 712–721.

Staring, M., Klein, S., & Pluim, J. P. W. (2006). Nonrigid registration using a rigidity constraint. In J. M. Reihnardt, & J. P. W. Pluim (Eds.), *Proceedings of the SPIE 2006: Medical imaging, 2006* (Vol. 6144, pp. 1–10). Boca Raton: SPIE.

Trottenberg, U., Oosterlee, C., & Schüller, A. (2001). *Multigrid*. London: Academic Press.

Wells III, W. M., Viola, P., Atsumi, H., Nakajima, S., & Kikinis, R. (1996). Multi-modal volume registration by maximization of mutual information. *Medical Image Analysis*, *1*(1), 35–51.

Yoo, T. S. (2004). *Insight into images: Principles and practice for segmentation, registration, and image analysis*. AK Peters Ltd.

Zitová, B., & Flusser, J. (2003). Image registration methods: a survey. *Image and Vision Computing*, *21*(11), 977–1000.