MULTIMODAL IMAGE REGISTRATION USING A VARIATIONAL APPROACH*

STEFAN HENN† AND KRISTIAN WITSCH†

Abstract. This paper presents an approach to obtain a deformation which matches two images acquired from different medical imaging modalities. This problem arises in the investigation of human brains. Two distance functionals for the images are proposed with different pros and cons. These functionals are to be minimized. We add a smoothing term to the minimization problem which retains certain desired elastic features in the solution. At each minimization step an approximate solution for the linearized problem is computed with a multigrid method as an inner iteration. Furthermore, we use a multiresolution minimization approach to obtain a suitable initial guess. Finally, we present some experimental results for registration problems of synthetic images and for a real computer tomography (CT)–magnetic resonance imaging (MRI) registration.

1. Introduction. Image registration, also known as image matching or image mapping, is a widely used method in medical image analysis, having applications in various domains; see, e.g., [1, 5, 8, 10, 12, 21]. A good survey of a part of the practical applications is given in [6, 14] and the references therein.

Given are two images, a reference R and a template T of the same object using the same or different imaging modalities. We assume that in continuous variables the images can be represented by compactly supported functions $T, R : \Omega \to G$, which map points of a domain $\Omega \subset \mathbb{R}^2$ to a domain $G \subset \mathbb{R}$. This means that the map associates with each pixel (picture element) $x = (x_1, x_2)^t \in \Omega \cap \mathbb{R}^2$ its intensities $T(x)$ and $R(x)$. For the purpose of numerical computation Ω will simply be the unit square $[0, 1]^2$ and $G = [0, 1]$ for gray-scale images. We assume that T is distorted by an invertible deformation ϕ^{-1}. We search for a transformation

$$\phi(u)(\cdot) \, : \, \mathbb{R}^2 \to \mathbb{R}^2, \quad \phi(u)(x) \, : \, x \mapsto x - u(x)$$

that depends on the unknown displacements $u = (u_1, u_2)^t : \mathbb{R}^2 \to \mathbb{R}^2$ (whose components are functions of the variables $x = (x_1, x_2)^t$). The goal of image registration is to determine u from the space of displacements $\mathcal{X}$ in such a way that the transformed template $T \circ \phi(u)$ matches the reference R. For a functional $\mathcal{D}[R, T, u]$, which measures the disparity between $T \circ \phi(u)$ and R, the image registration problem can be identified, in that way, with the following minimization problem:

(1.1)   Find $u \in \mathcal{X}$ such that $\mathcal{D}[R, T, u]$ is minimal.

In the situation where the intensities of the given images are comparable, a proper choice of $\mathcal{D}$ is the so-called sum of squared differences

$$\mathcal{D}^{SSD}[R, T, u] = \int_\Omega \left(T(x_1 - u_1(x_1, x_2), x_2 - u_2(x_1, x_2)) - R(x_1, x_2) \right)^2 dx_1 dx_2.$$

---

†Mathematisches Institut, Heinrich-Heine-Universität Düsseldorf, Universitätsstraße 1, D-40225 Düsseldorf, Germany (henn@am.uni-duesseldorf.de, witsch@am.uni-duesseldorf.de).

This is a common criterion. It is used, for example, in the case where the images are recorded with the same imaging machinery, the so-called monomodal image registration, or in the case of registration of MRI to histological images. In general, if the images are recorded with different imaging machinery, the so-called multimodal registration, the $\mathcal{D}^{SSD}$ functional is not an appropriate measure. The main reason is that the same structures may have quite different gray values in the multimodal case. Multimodal image registration is required, e.g., in radiotherapy planning where the representation of a tumor is much better performed using MRI but the present planning (e.g., dose calculation) has to be done based on CT.

The aim of this paper is to examine appropriate measurements for the general multimodal registration problem and to design the deformations in a suitable way to make certain that they have desired features. This leads to a variational problem similar to many other approaches, which have been investigated in the last decade for a variety of purposes in digital image processing [1, 2, 3, 7, 12, 16, 17, 18, 20], including image denoising, image segmentation, curve evolution, mathematical morphology, speech recognition, optical flow computation, and monomodal image matching.

The paper is organized as follows. In section 2, we will discuss suitable measures for the multimodal registration problem. We present two approaches using different distance metrics.

In section 3, we present an iterative method to minimize the distance functionals that can be regarded as an explicit discretization of a classical gradient flow method (e.g., see [9]) . Here, the basic idea is the computation of a smoothed approximation of a linearized problem. To this purpose, we add a smoothing term, which arises in the investigation in the theory of elastic deformations. The corresponding Euler–Lagrange equations are a coupled system of linear partial differential equations (PDEs). The discretized PDEs are solved by using multigrid methods with optimal complexity. Throughout the iteration we solve a one-dimensional minimization problem, which makes sure that in every iteration step the functional decreases. Finally, in section 4, we present some experimental results for a synthetic example and for a real CT–MRI registration.

2. Multimodal distance measures. The goal of this section is to introduce two different distance measures for the multimodal registration problem (1.1). The multimodal registration task can be illustrated by the example depicted in Figure 6. Here, some corresponding structures of the images have nearly equal gray values, while other structures have completely different (in the worst case, inverse) intensities for corresponding structures. For convenience, we will use in this section $\xi_1^u = T(x-u(x))$ for the intensity values of the deformed template and $\xi_2 = R(x_1, x_2)$ for the reference. We suppose ξ_2 and ξ_1^u to be random variables taking values in the set of images $\mathcal{Y}$, whose probability mass functions are given by d^T (resp., d^R). The joint distribution $d^{T,R}$ summarizes the co-occurrence of events from ξ_1^u and ξ_2 and describes "how random" the "joint variable" (ξ_1^u, ξ_2) is.

2.1. Distance between the probability density functions. Our first approach to solving the nonlinear matching problem is to map corresponding structures in the images with nearly equal probability. Therefore, we assume that the pixel density in the template image has a probability similar to the corresponding pixel density in the reference image.

This leads us to search for u so that $d^T(\xi_1^u) \approx d^R(\xi_2)$. We address the multimodal

image matching problem by minimizing the nonlinear functional

$$(2.1) \qquad \mathcal{D}^{DF}[R, T, u] = \int_{\Omega} \left(d^T(\xi_1^u) - d^R(\xi_2)\right)^2 dx_1 dx_2.$$

Of course, this approach imposes restrictions on the registration process, but it works well for many image modalities, which are interesting for medical applications, such as CT, MR, or histological image data. A minimizer u of (2.1) is characterized by the necessary condition $\frac{\partial}{\partial u} \mathcal{D}^{DF}(u) = 0$. This condition can be expressed by

$$\frac{\partial}{\partial u} \mathcal{D}^{DF}[R, T, u] = \int_{\Omega} 2 \cdot \left(d^T(\xi_1^u) - d^R(\xi_2)\right) \cdot \frac{\partial}{\partial u} d^T(\xi_1^u) \, dx_1 dx_2 = 0.$$

With the Jacobian J of a function we determine

$$\frac{\partial d^T(\xi_1^u)}{\partial u} d\Omega = J_{d^T}(T) \cdot \underbrace{J_T(\phi(u))}_{=\nabla T(\phi(u))} \cdot \underbrace{J_\phi(u)}_{=-I} = -J_{d^T}(T) \cdot \nabla T \circ \phi(u)$$

and get overall

$$(2.2) \qquad \tfrac{\partial}{\partial u} \mathcal{D}^{DF}[R, T, u] = -\int_{\Omega} 2 \cdot \left(d^T(\xi_1^u) - d^R(\xi_2)\right) \cdot J_{d^T}(\xi_1^u) \cdot \nabla T(\phi(u)) \, dx_1 dx_2 = 0.$$

2.2. The mutual information. An information theoretical approach for image registration was suggested by Viola, Wells, and others in [23, 22]. They search for an affine-linear transformation so that the mutual information (or transinformation) is maximized.

The mutual information is a functional of u and is defined by the distance between the joint distribution $d^{T,R}(\xi_1^u, \xi_2)$ and the product distribution $d^T(\xi_1^u) \cdot d^R(\xi_2)$ of the random variables ξ_2 and ξ_1^u:

$$\begin{aligned} \mathcal{D}^{MI}[R, T, u] &= KL(d^{T,R}(\xi_1^u, \xi_2), d^T(\xi_1^u) \cdot d^R(\xi_2)) \\ &= \int_{\Omega} d^{T,R}(\xi_1^u, \xi_2) \log\left(\frac{d^{T,R}(\xi_1^u, \xi_2)}{d^T(\xi_1^u) \cdot d^R(\xi_2)}\right) dx_1 dx_2 \end{aligned}$$

with the Kullback–Leibler distance $KL(p, q) = \int p(x) \log\frac{p(x)}{q(x)} dx$ (see, e.g., [4]) between two probability density functions p and q. The mutual information $\mathcal{D}^{MI}[R, T, u]$ is positive. Note that $\mathcal{D}^{MI}[R, T, u] = 0$ if and only if ξ_1^u and ξ_2 are independent, i.e., $d^{T,R}(\xi_1^u, \xi_2) = d^T(\xi_1^u) \cdot d^R(\xi_2)$.

This makes sense: if they are independent, then the random variable ξ_2 can tell us nothing about the random variable ξ_1^u. The mutual information is maximum if the images are matched. Therefore the mutual information is a measure of alignment between the images. This signifies that we have to maximize $\mathcal{D}^{MI}[R, T, u]$ or equivalently minimize $\mathcal{D}^{-MI}[R, T, u] := -\mathcal{D}^{MI}[R, T, u]$.

A minimizer u of the negative mutual information is characterized by the necessary condition $\frac{\partial}{\partial u} \mathcal{D}^{-MI}[R, T, u] = 0$. This condition can be expressed by the Euler–Lagrange equation

$$(2.3)$$
$$\int_{\Omega} \left(\left(\log\left(\frac{d^{T,R}(\xi_1^u, \xi_2)}{d^R(\xi_2) \cdot d^T(\xi_1^u)}\right) + 1\right) d_1^{T,R}(\xi_1^u, \xi_2) + \frac{d^{T,R}(\xi_1^u, \xi_2) d_u^T(\xi_1^u)}{d^R(\xi_2) \cdot d^T(\xi_1^u)}\right) dx_1 dx_2 = 0,$$

where d_u^T denotes the partial derivative of d^T with respect to u, and $d_1^{T,R}$ denotes the partial derivative of $d^{T,R}$ with respect to the first variable and satisfies

$$d_1^{T,R}(\xi_1^u, \xi_2) = J_{d^{T,R}}(\xi_1^u) \cdot J_T \circ \phi(u) \cdot J_\phi(u) = -J_{d^{T,R}}(\xi_1^u) \cdot \nabla T(\phi(u)).$$

3. Iterative minimization of the distance functional. We solve the multi-modal image matching problem by approximatively minimizing a distance functional $\mathcal{D}[R,T,u]$, where $\mathcal{D} \in \{\mathcal{D}^{DF}, \mathcal{D}^{-MI}\}$. The functional is nonlinear in u and contains in general multiple local and global minima. In order to find a local minimum of the image registration problem, we replace the distance functional by its first order Taylor expansion around a known deformation. To obtain a smooth path from an initial deformation u^* to a minimizer of $\mathcal{D}$, we add a convex energy functional based on gradients of the actual registration energy. The energy used can be regarded as a penalty for "elastic stresses" resulting from the deformation of the images.

The iterative process can be considered to be an explicit discretization of a classical gradient-flow method for minimizing the distance functional. In every iteration step the resulting Euler–Lagrange equation is a coupled system of linear PDEs. The solution of the discretized PDE is obtained using a multigrid method.

3.1. Minimization of the distance functional. To minimize $\mathcal{D}$ for a given current approximation $u^{(k)} = (u_1^{(k)}, u_2^{(k)})^t$, we search for an approximation $u^{(k+1)} = (u_1^{(k+1)}, u_2^{(k+1)})^t$ so that

$$\mathcal{D}[R,T,u^{(k+1)}] < \mathcal{D}[R,T,u^{(k)}].$$

The functional $\mathcal{D}[R,T,u^{(k+1)}]$ is replaced by its linearization around $u^{(k)}$,

$$\mathcal{D}[R,T,u^{(k+1)}] \approx \mathcal{D}[R,T,u^{(k)}] + \left\langle \frac{\partial}{\partial u}\mathcal{D}[R,T,u^{(k)}], \delta u^{(k)} \right\rangle_{L_2(\Omega)},$$

with $\delta u^{(k)} = u^{(k+1)} - u^{(k)}$. We add a bilinear form $B : \overline{\mathcal{X}} \times \overline{\mathcal{X}} \to \mathbb{R}$ to the linearized functional with desired elastic properties of the deformations (see section 3.2) and get the following minimization problem:

$$\arg\min_{\delta u^{(k)}} \left\{ \left\langle \frac{\partial}{\partial u}\mathcal{D}[R,T,u^{(k)}], \delta u^{(k)} \right\rangle + B[\delta u^{(k)}, \delta u^{(k)}] \right\}.$$

The Euler–Lagrange equation is given by the linear variational equation

$$\left\langle L[\delta u^{(k)}] + \frac{\partial}{\partial u}\mathcal{D}[R,T,u^{(k)}], \varphi \right\rangle = 0 \quad \text{for all } \varphi \in \overline{\mathcal{X}},$$

where L denotes the canonical linear mapping $B[\delta u^{(k)}, \varphi] = \langle L[\delta u^{(k)}], \varphi \rangle$ which maps the derivatives of $\mathcal{D}$ into a more regular space $\overline{\mathcal{X}} \subset \mathcal{X}$. Therefore, we minimize $\mathcal{D}$ by successively determining

$$u^{(k+1)} = u^{(k)} + \delta u^{(k)} \text{ for } k = 0, 1, \dots$$

with an initial guess $u^{(0)} = u^*$ and the solution $\delta u^{(k)}$ of the linear system

(3.1)
$$L[\delta u^{(k)}] = -\frac{\partial}{\partial u}\mathcal{D}[R,T,u^{(k)}].$$

Here, every iteration step can be regarded as an explicit time-step

$$u^{(k+1)} = u^{(k)} - \frac{1}{\delta t_k} L^{-1}\left[\frac{\partial}{\partial u}\mathcal{D}[R,T,u^{(k)}] \right]$$

of length $\delta t_k = 1$ in a classical gradient-flow method to minimize the distance measure $\mathcal{D}$,

$$\begin{cases} \frac{\partial u(x,t)}{\partial t} & = & -L^{-1}\left[\frac{\partial}{\partial u}\mathcal{D}[R,T,u(x,t)]\right] & \text{on} & \Omega \times (0,\infty), \\ u(x,0) & = & u^*(x) & \text{on} & \Omega, \end{cases}$$

and boundary conditions for u on $\partial\Omega \times (0,\infty)$ which depend on the bilinear form B used for L. Of course, in general the sequence $\{\mathcal{D}[R,T,u^{(k)}]\}_{k\in\mathbb{N}}$ is not decreasing necessarily for a constant time-step δt_k. This problem can be overcome by considering the function $\mathcal{D}[R,T,u^{(k)} + \tau_k d^{(k)}]$ for an artificial evolution parameter τ_k and $d^{(k)} := \delta u^{(k)}/||\delta u^{(k)}||_\infty$.

When the parameter τ_k is small enough, this is a decreasing function, due to the direction of $d^{(k)}$. We choose the parameter τ_k as a solution of the following one-dimensional minimization problem:

(3.2)   Find τ_k such that $\tau_k = \arg\min_{\tau\in[0,2]} \mathcal{D}[R,T,u^{(k)} + \tau \cdot d^{(k)}]$.

This means that the template is moved in one iteration step by at most two pixels. In practice, this seems to be a reasonable compromise between convergence speed and robustness. We stop the iteration when $\nabla\mathcal{D}[R,T,u^{(k)}] \approx 0$ and get the following algorithm.

ALGORITHM 3.1. *Iterative minimization of* $\mathcal{D}[R,T,u]$ $(\mathcal{D} \in \{\mathcal{D}^{-MI}, \mathcal{D}^{DF}\})$.
     $k = 0;\ u^{(0)} = 0;$
     repeat
       calculate $\frac{\partial}{\partial u}\mathcal{D}[R,T,u^{(k)}]$ *by* (2.2) *or* (2.3)
       compute $\delta u^{(k)}$ *from* (3.1)
       set $d^{(k)} = \delta u^{(k)}/||\delta u^{(k)}||_\infty$
       compute τ_k *by solving problem* (3.2)
       set $u^{(k+1)} = u^{(k)} + \tau_k \cdot d^{(k)}$
       set $k = k + 1$
     until $||\nabla\mathcal{D}[R,T,u^{(k+1)}]||^2 \leq 10^{-8}$

3.2. Elastic registration. The monomodal image registration approach proposed in [11] leads us to choose a smoothing term of the form

$$B^{el}[u,v]$$

$$= \int_\Omega \left[2\mu \sum_{i,j=1}^2 \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) \cdot \left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i}\right) + \lambda \cdot \left(\sum_{i=1}^2 \frac{\partial u_i}{\partial x_i}\right) \cdot \left(\sum_{i=1}^2 \frac{\partial v_i}{\partial x_i}\right)\right] d\Omega.$$

The so-called Lamé constants $\lambda \geq 0$ and $\mu > 0$ reflect material properties. They are chosen so that the changes in volume are maximal (i.e., $\lambda = 0$ and $\mu = \frac{1}{2}$). In this case, the linear differential operator L^{el} is given by

$$L^{el}[u_1(x_1,x_2), u_2(x_1,x_2)] = \begin{cases} -\Delta u_1(x_1,x_2) - (\partial_{x_1,x_1}u_1(x_1,x_2) + \partial_{x_1,x_2}u_2(x_1,x_2)) \\ -\Delta u_2(x_1,x_2) - (\partial_{x_1,x_2}u_1(x_1,x_2) + \partial_{x_2,x_2}u_2(x_1,x_2)) \end{cases}$$

for $(x_1,x_2) \in \Omega$, the displacements $u = (u_1(x_1,x_2), u_2(x_1,x_2))^t)$, and the Laplace operator $\Delta := \partial_{x_1,x_1} + \partial_{x_2,x_2}$. In medical applications, it is appropriate to use $\overline{\mathcal{X}} = H_0^1(\Omega) \times H_0^1(\Omega)$, which yields Dirichlet boundary conditions

$$u_l(x_1,x_2) = 0 \quad \text{for} \quad (x_1,x_2) \in \partial\Omega \quad \text{and} \quad l = 1,2.$$

Here, the bilinear form models the brain data like an elastic medium, measures the energy of the elastic deformation, is isotropic in the directions, and is neutral with respect to translations and rotations but penalizes these transformations by the Dirichlet boundary conditions.

3.3. Numerical solution. Solving the PDE (3.1) is the time-consuming part of the minimization process. In the following, we describe briefly the discretization, approximation, and the solution method for solving the PDE numerically.

3.3.1. Discretization. In practice the images T and R are encoded as two-dimensional arrays of image intensities denoted by T_h (resp., R_h). For the purpose of the numerical computation of the deformation u, we choose a finite difference discretization. With the regular grid

$$G_h = \left\{ x \in \mathbb{R}^2 : x = \left(\frac{h}{2} + ih, \frac{h}{2} + jh \right), \quad i,j = 0,1,\dots,\infty \right\}$$

we arrange the grid points of $\Omega_h = G_h \cap \Omega$ at the centers of the two-dimensional image array (pixel-centered discretization). The discrete displacement vector is given by

$$u_h(x_h) = (u_{1,h}, u_{2,h})^t \in \mathcal{F}(\Omega_h) \times \mathcal{F}(\Omega_h)$$

with a grid function $\mathcal{F}$ defined by $\mathcal{F} : \Omega_h \to \Omega_h$.

3.3.2. Approximation. We approximate the partial derivatives by second order approximations. This yields the following second order stencils:

$$-\Delta u = \frac{1}{h^2} \big(4 \cdot u(x_1, x_2) - u(x_1 + h, x_2) - u(x_1 - h, x_2) - u(x_1, x_2 + h)$$
$$- u(x_1, x_2 - h) \big) + \mathcal{O}(h^2)$$

and

$$u_{x_1 x_2} = \frac{1}{4h^2} \big(u(x_1 + h, x_2 - h) - u(x_1 - h, x_2 - h) - u(x_1 + h, x_2 + h)$$
$$+ u(x_1 - h, x_2 + h) \big) + \mathcal{O}(h^2).$$

With $f_h := -\frac{\partial}{\partial u} \mathcal{D}[R_h, T_h, u_h^{(k)}]$ this leads to the discrete difference operator

$$L_h^{el}[u_h] = \begin{cases} \frac{1}{h^2} \begin{bmatrix} 0 & -1 & 0 \\ -2 & 6 & -2 \\ 0 & -1 & 0 \end{bmatrix} u_{1,h} + \frac{1}{4h^2} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix} u_{2,h} & = & f_{1,h} \text{ on } \Omega_h, \\ \\ \frac{1}{h^2} \begin{bmatrix} 0 & -2 & 0 \\ -1 & 6 & -1 \\ 0 & -2 & 0 \end{bmatrix} u_{2,h} + \frac{1}{4h^2} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix} u_{1,h} & = & f_{2,h} \text{ on } \Omega_h, \\ \\ u_{1,h} = u_{2,h} & = & 0 \text{ on } \partial\Omega_h. \end{cases}$$

In two-dimensional applications with typically $2^8 \times 2^8$ picture elements, the linear system with unknowns

$$u_h = (u_1^{(1,1)}, \dots, u_1^{(n-1,n-1)}, u_2^{(1,1)}, \dots, u_2^{(n-1,n-1)}) \in \mathcal{F}(\Omega_h) \times \mathcal{F}(\Omega_h)$$

and right-hand side

$$f_h = (f_1^{(1,1)}, \dots, f_1^{(n-1,n-1)}, f_2^{(1,1)}, \dots, f_2^{(n-1,n-1)}) \in \mathcal{F}(\Omega_h) \times \mathcal{F}(\Omega_h)$$

includes up to 10^5 equations and unknowns.

3.3.3. Multigrid method. The PDE is solved by a classical multigrid correction scheme with one relaxation before and one after the coarse grid correction for smoothing the error by a V-cycle. On the coarsest grid, which includes only one inner grid point, we solve the defect equation directly. We use within the correction scheme a full weighting restriction for defect and bilinear interpolation for the corrections. Some different classical relaxation methods such as Gauss–Seidel, Jacobi, and variations thereof were implemented. It turns out that a Gauss–Seidel relaxation with red-black ordering of the grid points and an over-relaxation parameter $\omega = 1.3$ leads to the best convergence rates for the multigrid correction scheme. This relaxation is also used for the following results.

3.3.4. Numerical results. The convergence rates of our multigrid implementation are very good and independent of the grid size h. In order to show this we solve the linear system $L^{el}[u] = f$ with given right-hand side

$$f_1(x_1, x_2) = \pi^2 \cdot (4 \cdot \sin 2\pi x_2 \cdot (2 \cdot \cos 2\pi x_1 - 1) - \cos \pi(x_1 + x_2) + \sin \pi x_1 \cdot \sin \pi x_2),$$
$$f_2(x_1, x_2) = \pi^2 \cdot (4 \cdot \sin 2\pi x_2 \cdot (1 - 2 \cdot \cos 2\pi x_1) - \cos \pi(x_1 + x_2) + \sin \pi x_1 \cdot \sin \pi x_2)$$

and corresponding exact solution

$$u_1^{ex}(x_1, x_2) = \sin 2\pi x_2 \cdot (\cos 2\pi x_1 - 1) + \sin \pi x_1 \cdot \sin \pi x_2,$$
$$u_2^{ex}(x_1, x_2) = \sin 2\pi x_1 \cdot (1 - \cos 2\pi x_2) + \sin \pi x_1 \cdot \sin \pi x_2$$

for different mesh sizes h_l. For this example, we present in Table 3.1 the errors $||I_h u_1^{ex} - u_{1,h}^{(\nu)}||_2$ for the numerical solution of the first component $u_{1,h}^{(\nu)}$ after the νth multigrid iteration. It can be seen that the multigrid iteration converges up to an accuracy of 10^{-08} after the seventh iteration—independent of the grid size h_l. Furthermore, the error is reduced after the first iteration to approximately $3.2 \cdot 10^{-02}$. This indicates optimal multigrid convergence.

Note that the second order approximation can be observed by the discretization errors, which decrease quadratically.

TABLE 3.1
Error development for a V-cycle.

	$h_l = \frac{1}{16}$	$h_l = \frac{1}{32}$	$h_l = \frac{1}{64}$	$h_l = \frac{1}{128}$	$h_l = \frac{1}{256}$
$\nu = 1$	$2.9980 \cdot 10^{-02}$	$3.218 \cdot 10^{-02}$	$3.27 \cdot 10^{-02}$	$3.28 \cdot 10^{-02}$	$3.2 \cdot 10^{-02}$
$\nu = 2$	$1.9861 \cdot 10^{-03}$	$2.777 \cdot 10^{-03}$	$3.05 \cdot 10^{-03}$	$3.12 \cdot 10^{-03}$	$3.1 \cdot 10^{-03}$
$\nu = 3$	$1.4260 \cdot 10^{-03}$	$3.733 \cdot 10^{-04}$	$3.55 \cdot 10^{-04}$	$3.89 \cdot 10^{-04}$	$3.9 \cdot 10^{-04}$
$\nu = 4$	$1.5455 \cdot 10^{-03}$	$3.713 \cdot 10^{-04}$	$9.20 \cdot 10^{-05}$	$5.51 \cdot 10^{-05}$	$5.8 \cdot 10^{-05}$
$\nu = 5$	$1.5580 \cdot 10^{-03}$	$3.862 \cdot 10^{-04}$	$9.47 \cdot 10^{-05}$	$2.33 \cdot 10^{-05}$	$9.8 \cdot 10^{-06}$
$\nu = 6$	$1.5593 \cdot 10^{-03}$	$3.882 \cdot 10^{-04}$	$9.67 \cdot 10^{-05}$	$2.39 \cdot 10^{-05}$	$5.9 \cdot 10^{-06}$
$\nu = 7$	$1.5594 \cdot 10^{-03}$	$3.885 \cdot 10^{-04}$	$9.70 \cdot 10^{-05}$	$2.42 \cdot 10^{-05}$	$6.0 \cdot 10^{-06}$
$\nu = 8$	$1.5594 \cdot 10^{-03}$	$3.885 \cdot 10^{-04}$	$9.70 \cdot 10^{-05}$	$2.42 \cdot 10^{-05}$	$6.0 \cdot 10^{-06}$

3.4. Multiresolution minimization. The distance functionals $\mathcal{D}^{-MI}$ and $\mathcal{D}^{DF}$ are nonlinear and may have many global and local minima. In section 3.2 we have introduced an algorithm which finds a local minimum close to an initial guess u^*. Without a priori information, it is suitable to start the image registration with $u^{(0)} = 0$, which corresponds to the identity map $\phi(x) = x$. For a given small pixel width h (resp., the step size of the discretization of the linear PDE (3.1)) this leads to a deformation that mainly adapts structures with size h and neglects larger structures.

Here, the minimization process sometimes finds only a poor local minimum and requires many time-consuming iteration steps. Both disadvantages can be avoided by minimizing the distance functional $\mathcal{D}[R, T, u]$ at first on coarser resolutions.

We approximate the sets $\mathcal{X}_h$ and $\mathcal{Y}_h$ by a sequence of subsets

$$\mathcal{X}_{h_L} \subset \mathcal{X}_{h_{L-1}} \subset \cdots \subset \mathcal{X}_h \subset \mathcal{X} \quad \text{and} \quad \mathcal{Y}_{h_L} \subset \mathcal{Y}_{h_{L-1}} \subset \cdots \subset \mathcal{Y}_h \subset \mathcal{Y}$$

with increasing mesh size $h_l = 2^l h$ for $l = 0, 1, \ldots, L$ and decreasing dimension. We get a sequence of minimization problems

$$(M_{h_l}) \qquad u_{h_l}^* = \arg \min_{u_{h_l} \in \mathcal{X}_{h_l}} \left\{ \mathcal{D}[R_{h_l}, T_{h_l}, u_{h_l}] \right\}$$

defined on the sets $\mathcal{X}_{h_l}$. With a bilinear interpolation operator $I_l^{l-1} : \mathcal{X}_{2^l h} \to \mathcal{X}_{2^{(l-1)} h}$ we transform the solution $u_{h_l}^* \in \mathcal{X}_{h_l}$ of M_{h_l} onto the next finer resolution. Then $I_l^{l-1}(u_{h_l}^*) \in \mathcal{X}_{2^{l-1} h}$ is a suitable initial guess for the minimization of $M_{h_{l-1}}$. The multiresolution minimization algorithm can be described as follows.

ALGORITHM 3.2. *Multiresolution minimization of $\mathcal{D}[R, T, u]$.*

> $l = L$; $u_{h_L} = 0$;
> *repeat*
>   *if ($l == L$)*
>   *find a local minimum $u_{h_L}^*$ of M_{h_L}, starting from u_{h_L}*
>   *else*
>   *interpolate the local minimum $u_{h_l} = I_{l+1}^l(u_{h_{l+1}}^*)$*
>   *find a local minimum $u_{h_l}^*$ of M_{h_l}, starting from u_{h_l}*
>   *endif*
>   *set $l = l - 1$*
> *until ($l == 0$)*

4. Results. To demonstrate the efficiency of the proposed matching algorithm, we present numerical results on two experiments of varying difficulty. Throughout this section we denote the minimization of the density-based functional in (2.1) as approach 1 and the mutual information in (2.2) as approach 2.

4.1. Specifics of the solution method. To solve the PDE in (3.1), we used in all examples a multigrid correction scheme for smoothing the error with the maximal number of coarse grids, one relaxation before, and one after the coarse grid correction, with full weighting restriction and bilinear interpolation. For all examples, we use within the multiresolution minimization (Algorithm 3.2) the injection operator for the restriction of the image data and the bilinear interpolation to transfer the initial guess on the next finer grid. During the iteration the probability functions are estimated by the approach introduced in [23]. The results presented in Figure 2 (resp., Figure 7) are calculated by using a two-level strategy (i.e., $L = 1$ in Algorithm 3.2).

4.2. Experiments with synthetic images. In order to show the working principle, we present the registration of two synthetic images (256×256) displayed in Figure 1. The images contain simulated structures of varying size which are delineated in different modalities. The synthetic images should be matched to give an indication of the behavior of the two proposed approaches. Figure 2 displays the calculated results. Both approaches match the different structures onto the correspondent reference structures. The checkerboard views of the results show smooth transitions between the structures.
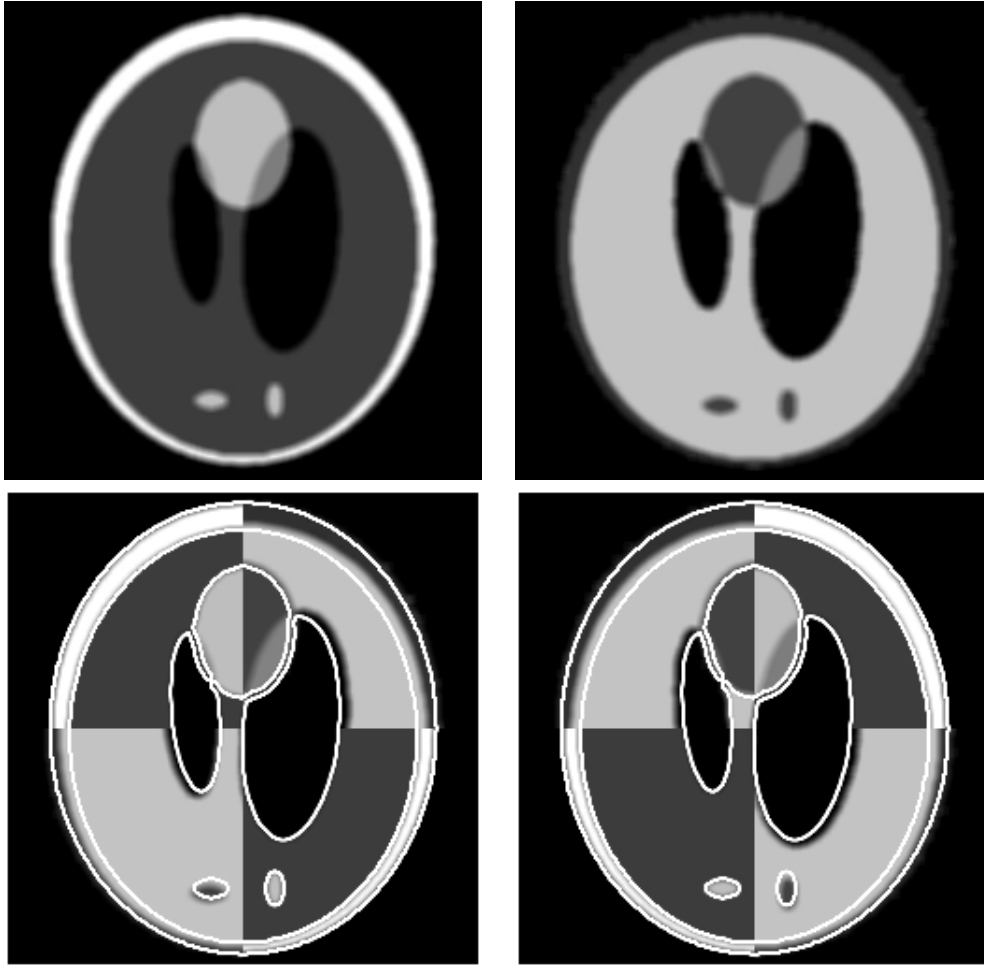
FIG. 1. *Top, left to right: Reference image and template image. Bottom, left to right: Template and reference image mixed like a 2×2 checkerboard with superimposed reference contour.*

These findings can be stressed by the joint probability distribution $d^{T,R}$ of the two images, which can be graphically displayed as a $2D$-histogram. The $2D$-histogram has an entry at position (t, r) if there exists at least one pixel with value r in the reference image and a value t in the template image. The pictures in Figure 3 are made by the MATLAB *spy* command and display the sparsity pattern of the $2D$-histogram. The $2D$-histogram of the images before registration is displayed at the top of Figure 3. The misalignment of the images can be illustrated by the many diffuse clusters in the $2D$-histogram.

The $2D$-histograms after registration via approaches 1 and 2 consist of three central peaks. The peak around $(0, 0)$ corresponds to the background, the large ellipsoid structure corresponds to the peak around $(60, 190)$, and the ellipsoid-shaped outer structure corresponds to the peak around $(255, 50)$. But there are subtle distinctions between the $2D$-histograms. The result obtained by approach 1 includes a high profile cluster from $(0, 0)$ to $(150, 150)$ in the associated $2D$-histogram. This corresponds to aligned structures with similar gray values. By matching the images by the second approach this correlation is not so pronounced in the resulting $2D$-histogram. Another
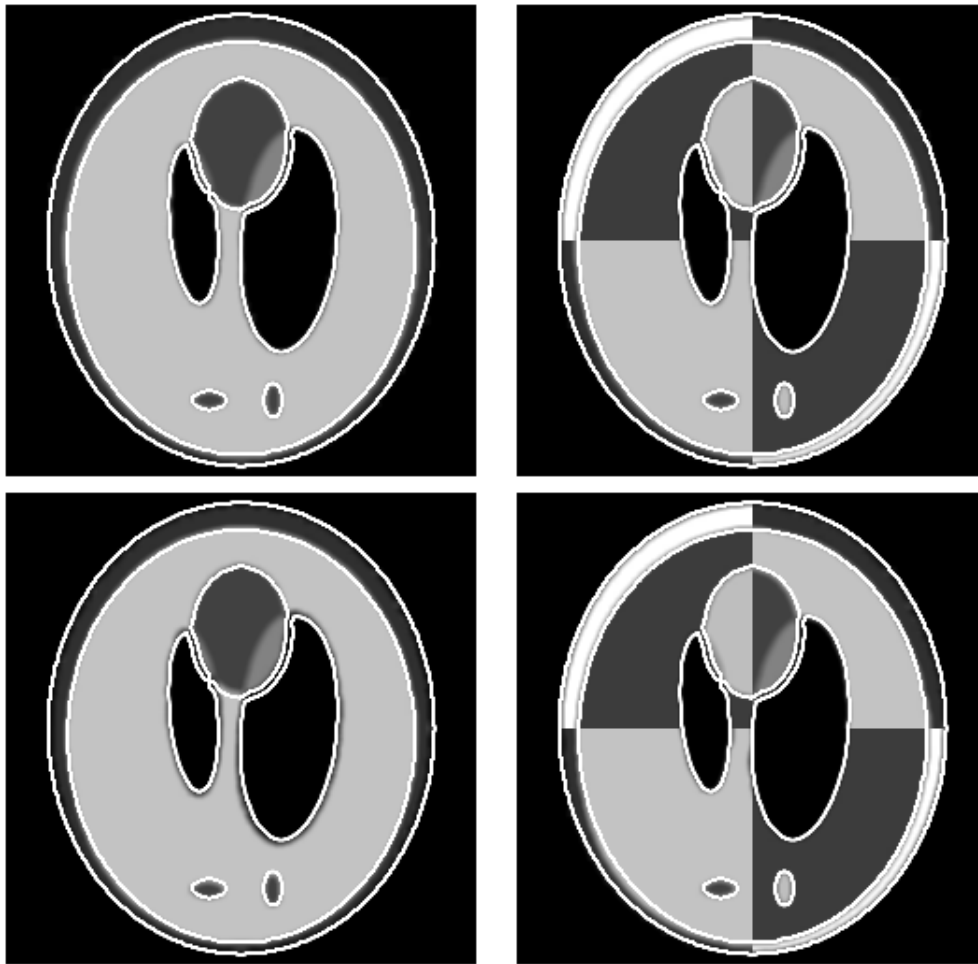
Fig. 2. *Top, left to right: Template deformed by the solution of approach 1 and superimposed reference contour, template deformed by the solution of approach 1 and reference image mixed like a 2×2 checkerboard with superimposed reference contour. Bottom, left to right: Template deformed by the solution of approach 2 and superimposed reference contour, template deformed by the solution of approach 2 and reference image mixed like a 2×2 checkerboard with superimposed reference contour.*

difference is that for the first approach a black pixel of the template corresponds, as expected, only to a black pixel of the reference. For the second approach some of the template pixels with gray values between 1 and 100 are mapped onto black reference pixels; this is a subtle hint of misalignment, which could not be found by considering only the deformed templates in Figure 2.

The graphs in Figures 4 and 5 display the decreasing distance measures $\mathcal{D}$ between the images after each minimization step of both approaches. The graphs are scaled by a work unit (WU), which corresponds to the effort of a multigrid correction scheme on the finest grid. The approaches have been carried out two times. First, we minimize the functional on only one image resolution and, second, we determine an initial guess on a coarse grid of size 128×128. It can be seen that both strategies converge to the same local minimum. By comparing the WUs needed for both strategies, we can observe that the two-level strategy needs only 52% of the costs of the one-level strategy by using the first approach and 36% by using the second approach.
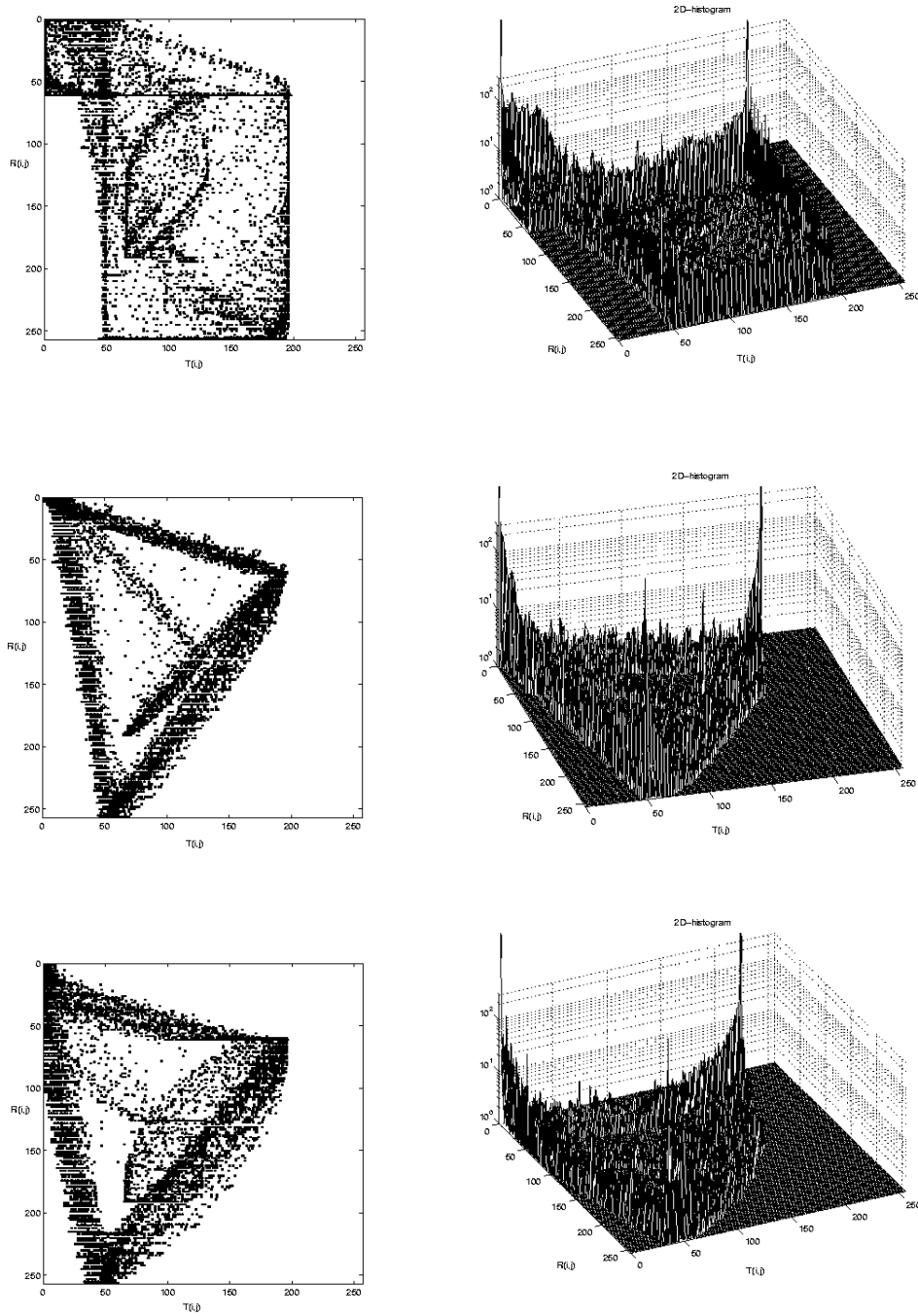
FIG. 3. *From left to right: Sparsity pattern of the 2D-histogram and the 2D-histogram. Top: 2D-histogram created from the template (x-axis) and the reference (y-axis). Middle: 2D-histogram created from the template deformed by the solution of approach 1 (x-axis) and the reference (y-axis). Bottom: 2D-histogram created from the template deformed by the solution of approach 2 (x-axis) and the reference (y-axis).*
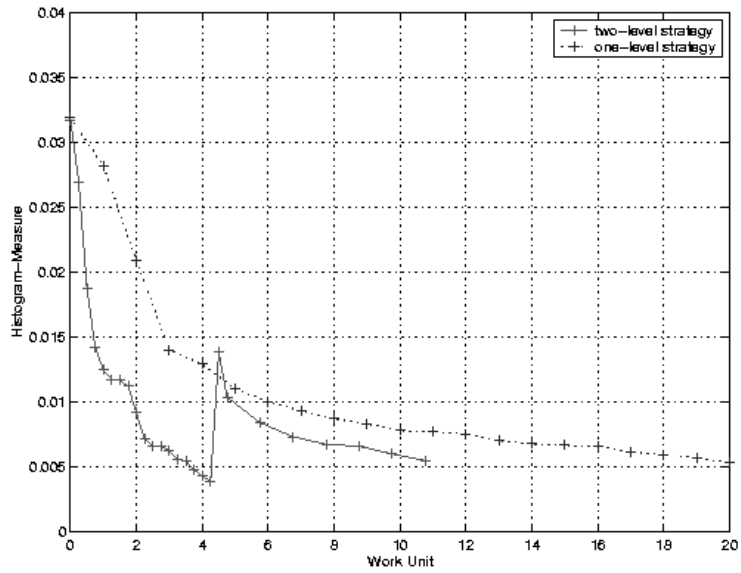
Fig. 4. *Distance measure $\mathcal{D}^{DF}[R, T, u^{(k)}]$ trace for the example displayed in Figure* 1.
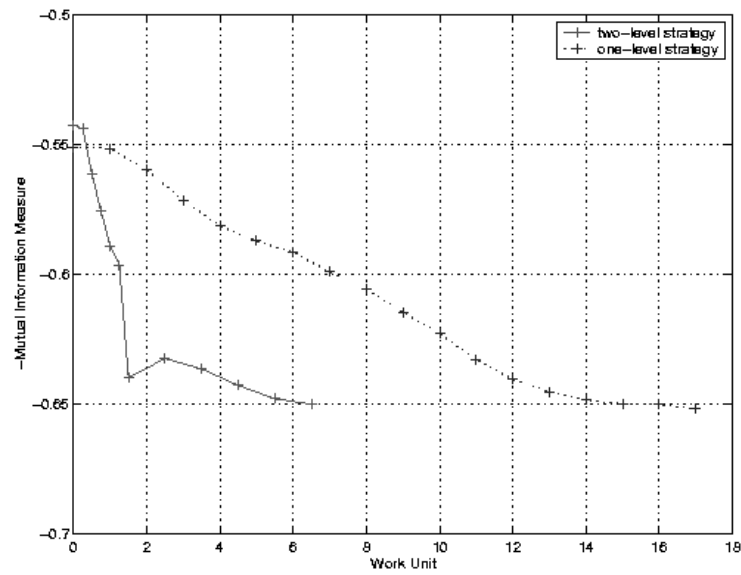

Fig. 5. *Distance measure $\mathcal{D}^{-MI}[R, T, u^{(k)}]$ trace for the example displayed in Figure* 1.
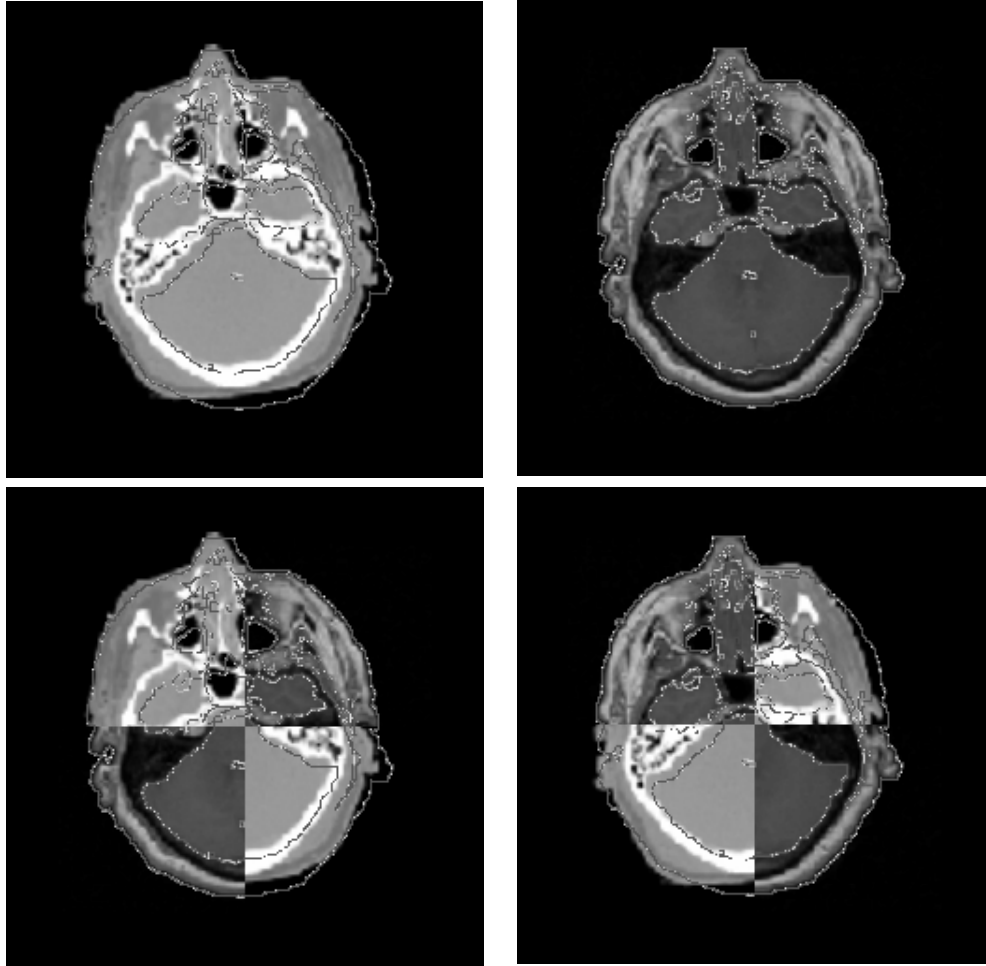
FIG. 6. *Top, left to right: CT template image, MR reference image. Bottom, left to right: Template image and reference image mixed like a 2×2 checkerboard. All images are presented with superimposed reference contour.*

4.3. A CT-MRI registration. In the next experiment, we present the registration of two images recorded by different image machineries. Figure 6 displays a CT slice (256×256) of a human brain and a corresponding MRI slice (256×256).

By using the two-level strategy, both approaches seem to match the main anatomical structures correctly. Here, the results are displayed in Figure 7. The checkerboard views show smooth transitions between the deformed structures of the CT image and the corresponding structures of the MR image.

The $2D$-histogram displayed in Figure 8 shows the typical broad ridge of MR soft tissue values corresponding to a narrow range of CT soft tissue values. An indicator for a good CT-MRI registration is that the high CT values representing the skull do not overlap higher MRI values corresponding to soft tissue. This criterion verifies the misregistration of the undeformed template and the reference image. After registration this effect is eliminated (see middle and bottom of Figure 8).
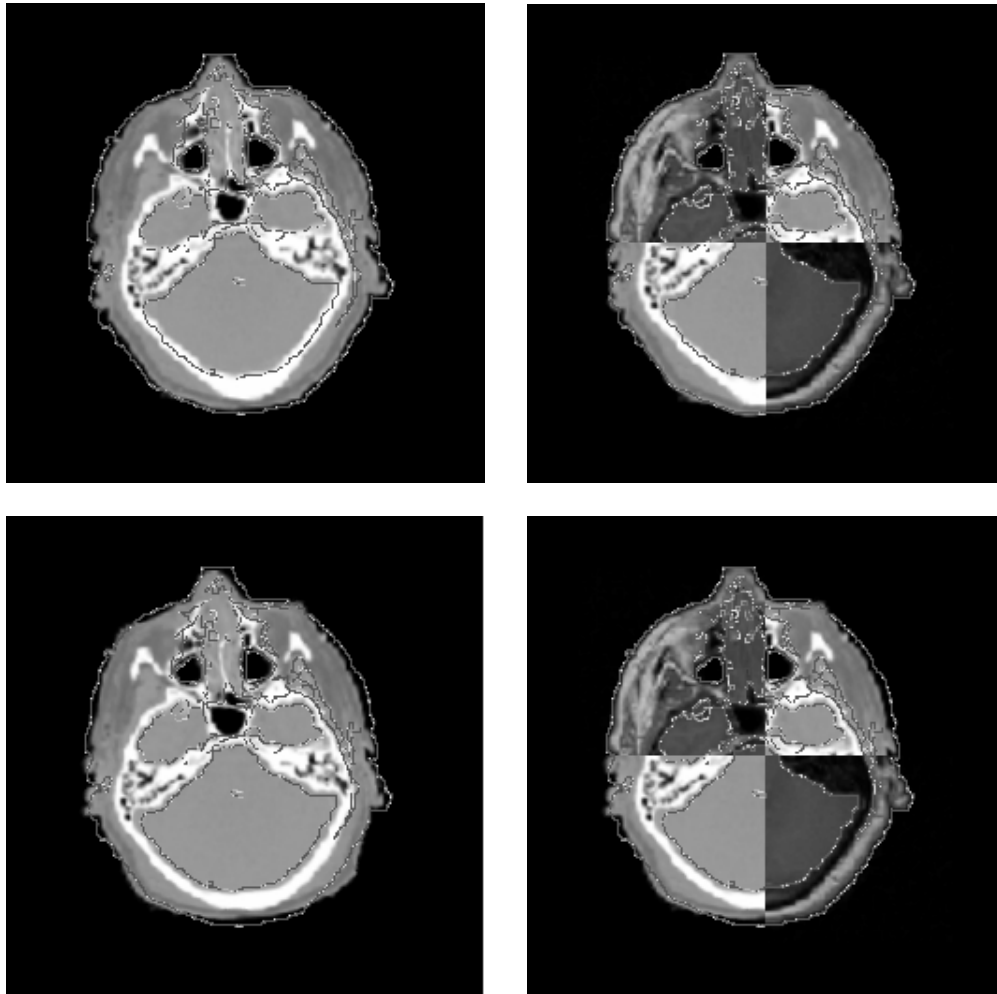
FIG. 7. *Top, left to right: CT template deformed by the solution of approach* 1 *in* (2.1) *and superimposed reference contour, CT template deformed by the solution of approach* 1 *and MR reference image mixed like a* 2×2 *checkerboard with superimposed reference contour. Bottom, left to right: CT template deformed by the solution of approach* 2 *in* (2.2) *and superimposed reference contour, CT template deformed by the solution of approach* 2 *and MR reference image mixed like a* 2×2 *checkerboard with superimposed reference contour.*

As for the synthetic experiment, approach 2 (in contrast to approach 1) maps black pixels of the template (CT image) onto pixels with gray values between 1 and 100 of the reference (MR image). In practice it is a difficult task, even for technicians, to evaluate the results on the basis of images or $2D$-histograms, but as shown it is possible that different deformations can lead to sensible alignments.

For the performance of the approaches we refer to Figures 9–10. The graph in Figure 9 displays the decreasing distance measures $\mathcal{D}^{DF}$. In this case, it can be seen that the one- and two-level strategies converge to the same local minimum. Note that the value of the distance functional for the two-level strategy drops rapidly on the coarse grid so that the two-level strategy needs only 50% of the costs of the one-level

strategy. The graph in Figure 10 displays the decreasing distance measures $\mathcal{D}^{-MI}$ for both a one-level and a two-level strategy. Here, the result (shown in Figure 11) of the one-level strategy is slightly worse than what is obtained with the two-level strategy (shown in Figure 7) in approximately the same amount of WUs. The reason is that, in this example, the one-level registration strategy mainly adapts small structures (e.g., the eyes) and neglects coarser structures (e.g., the skull).

5. Summary, conclusion, and future work. In this paper we have presented a pixel-based approach for nonlinear multimodal image matching. The main contribution of the paper is twofold. First, we measure the distance between the images by two different nonlinear functionals. Second, we achieve the minimization of these functionals by an iterative method which closely relates to a classical gradient-flow method with explicit discretization. We use a multilevel minimization to speed up the minimization process and to avoid irrelevant local minima. We model the images like an elastic medium by adding a bilinear form with desired elastic properties to the resulting Euler–Lagrange equations. We solve the resulting Euler–Lagrange equations by a multigrid method with optimal multigrid complexity $\mathcal{O}(N)$, with N the number of picture elements. For images consisting of 256×256 picture elements, our algorithm written in C takes less than one minute on a 400 MHz Linux-PC.

In the meantime there have been various research groups who investigated registration algorithms for multimodal images. Viola, Wells, and others [23, 22] introduced the mutual information as distance metric for medical images to obtain a global alignment of multimodal images. For nonlinear image matching, some authors [13] tried to apply some criterion developed for rigid matching using block matching techniques. The approach proposed in [19] assumes that, at the registration position, one image could be approximated in terms of the other by applying some intensity function. In [15] a mutual information metric is used to refine the position of control points in a thin plate spline-based registration.

Of course, the presented approach could not replace these techniques, but by considering more extensive structures (like the human brain) the presented approach will improve the registration process. On the other hand, the presented method is restricted to small nonlinear deformations; therefore in many cases a preregistration is necessary.
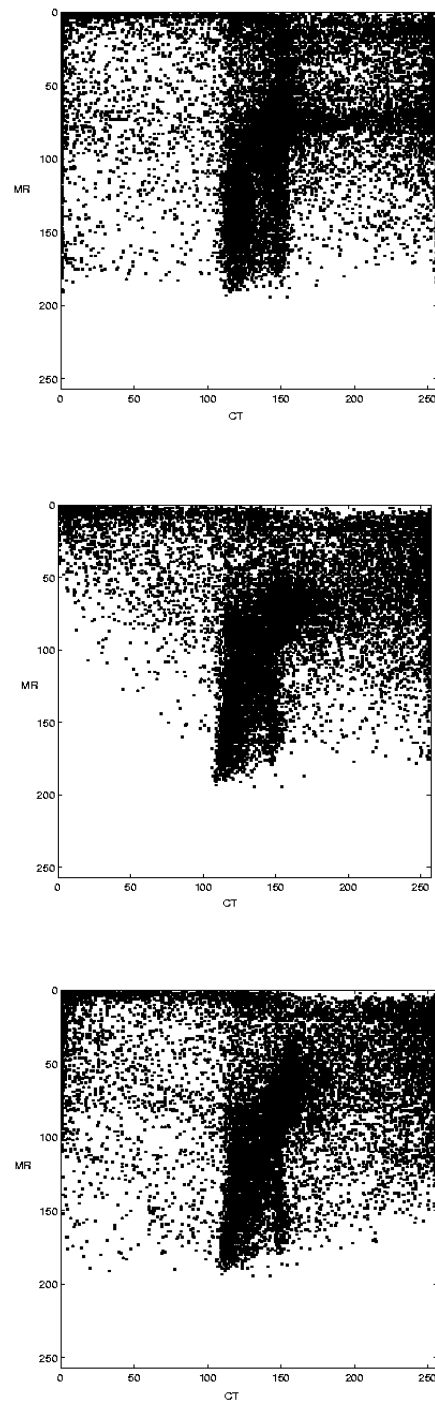
Fig. 8. *Top: 2D-histogram created from the CT template slice (x-axis) and the MRI reference slice (y-axis). Middle: 2D-histogram created from Algorithm* 3.1 *(x-axis) and the MRI reference slice (y-axis). Bottom: 2D-histogram created from Algorithm* 3.2 *(x-axis) and the MRI reference slice (y-axis).*
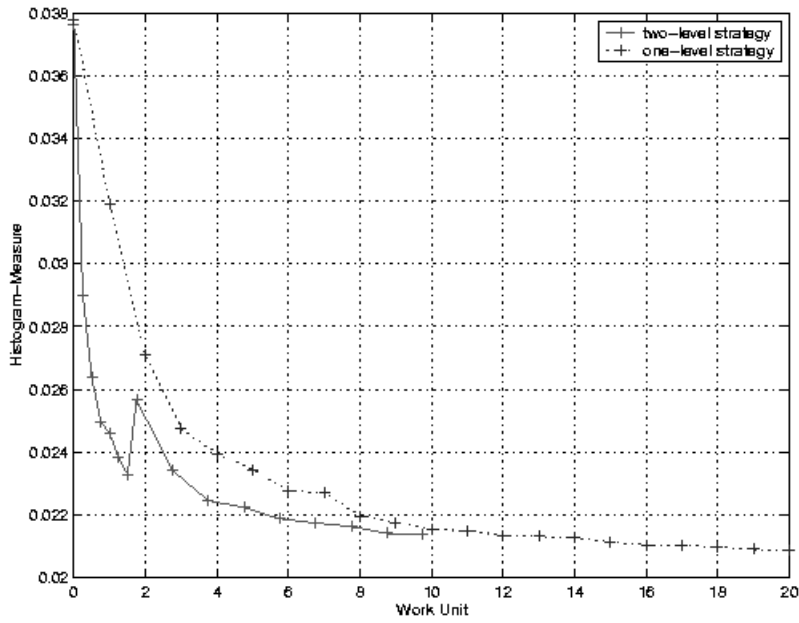
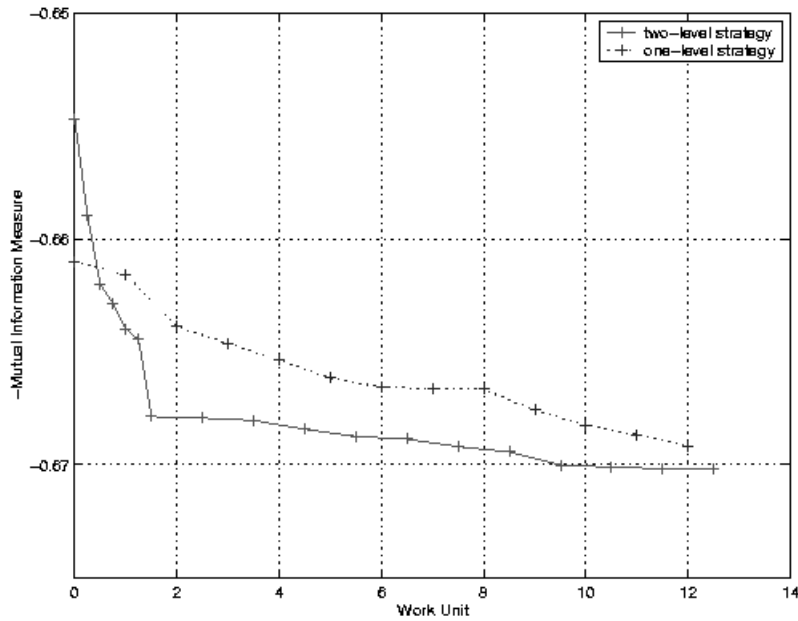FIG. 9. *Distance measure $\mathcal{D}^{DF}[R,T,u^{(k)}]$ trace for the example displayed in Figure* 6.

FIG. 10. *Distance measure $\mathcal{D}^{-MI}[R,T,u^{(k)}]$ trace for the example displayed in Figure* 6.
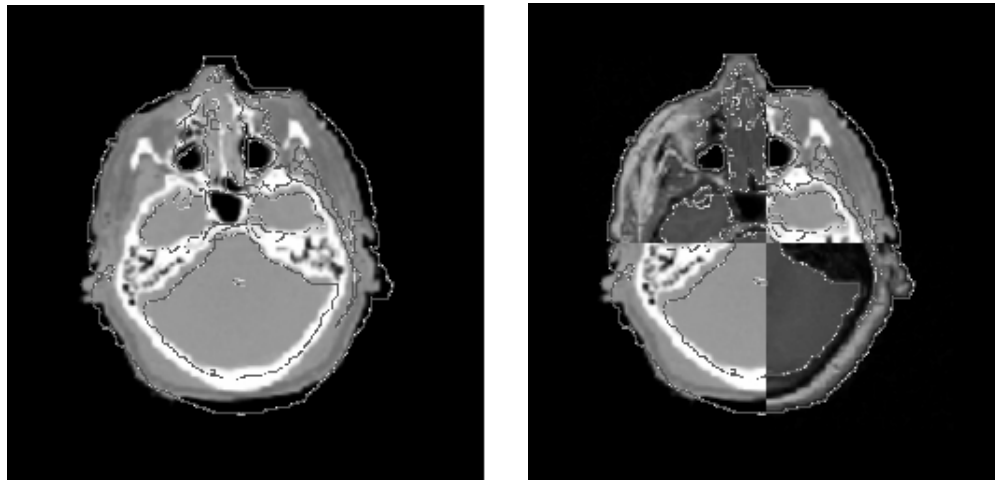
FIG. 11. *From left to right: CT template deformed by the solution of approach* 2 *in* (2.2) *and superimposed reference contour. CT template deformed by the solution of approach* 2 *and MR reference image mixed like a* 2 × 2 *checkerboard with superimposed reference contour.*

The experimental results indicate that both approaches proposed in this paper lead to satisfactory results. As shown in section 4, they seem to match the main anatomical structures for a CT-MRI registration problem. Consequently, both measures provide a meaningful indication of misalignment for the registration problems. We have discussed in section 4 the behavior of the algorithms with respect to the underlying measures and the multiresolution minimization strategy. It turns out that the numerical effort is about the same for both methods.

The similarity measure based on mutual information is highly nonlinear in u and contains many local and global minima caused by good local matches between some smaller substructures. As a consequence, the minimization process may lead to a "useless" local minimum. Typically, in this case a multiresolution minimization approach becomes necessary to find a meaningful minimum; see, for instance, the second example (Figure 11). This fact and the implementation of (2.3) lead to a significant programming overhead in comparison to the approach based on the density functional. In many cases the minimization also works here by using only the finest resolution. On the other hand, for this functional it is easy to construct examples where the registration fails, but in practical examples it often works surprisingly well. The computations are simpler than in the mutual information based approach, and therefore it seems useful to develop both methods. So, it is possible to choose the best method for application classes.

Based on our experience with unimodal image matching, the extension to a three-dimensional approach is straightforward and is planned for the near future.

REFERENCES

[1] Y. AMIT, *A nonlinear variational problem for image matching*, SIAM J. Sci. Comput., 15 (1994), pp. 207–224.
[2] G. AUBERT, R. DERICHE, AND P. KORNPROBST, *Computing optical flow via variational techniques*, SIAM J. Appl. Math., 60 (1999), pp. 156–182.

[3] G. Aubert and L. Vese, *A variational method in image recovery*, SIAM J. Numer. Anal., 34 (1997), pp. 1948–1979.

[4] R. Bahadur, *Some Limit Theorems in Statistics*, CBMS-NSF Reg. Conf. Ser. Appl. Math. 4, SIAM, Philadelphia, 1971.

[5] R. Bajcsy and S. Kovacic, *Multiresolution elastic matching*, Comput. Vision Graphics Image Process., 46 (1989), pp. 1–21.

[6] A. Barry, *Seeking signs of intelligence in the theory of control*, SIAM News, 30 (1997), p. 13.

[7] T.F. Chan, G.H. Golub, and P. Mulet, *A nonlinear primal-dual method for total variation-based image restoration*, SIAM J. Sci. Comput., 20 (1999), pp. 1964–1977.

[8] G. Christensen, M. Miller, M. Vannier, and U. Grenander, *Individualizing neuroanatomical atlases using a massively parallel computer*, IEEE Computer, 29 (1996), pp. 32–38.

[9] R. Courant, *Calculus of Variations*, Courant Institute of Mathematical Sciences, New York, NY, 1962.

[10] M. Davis, A. Khotanzad, D. Flaming, and S. Harms, *A physics based coordinate transformation for 3d medical images*, IEEE Trans. Medical Imaging, 16 (1997), pp. 317–328.

[11] S. Henn and K. Witsch, *A multigrid-approach for minimizing a nonlinear functional for digital image matching*, Computing, 64 (1999), pp. 339–348.

[12] S. Henn and K. Witsch, *Iterative multigrid regularization techniques for image matching*, SIAM J. Sci. Comput., 23 (2001), pp. 1077–1093.

[13] J. Maintz, E. Meijering, and M. Viergever, *General multimodal elastic registration based on mutual information*, in Medical Imaging 1998: Image Processing, SPIE Proc. 3338, K.M. Hanson, ed., SPIE, Bellingham, WA, pp. 144–154.

[14] J. Maintz and M. Viergever, *A survey of medical image registration*, Medical Image Analysis, 2 (1998), pp. 1–36.

[15] C. Meyer, J. Boes, B. Kim, P. Bland, K. Zasadny, P. Kison, K. Koral, K. Frey, and R. Wahl, *Demonstration of accuracy and clinical versatility of mutual information for automatic multimodality image fusion using affine and thin-plate spline warped geometric deformations*, Medical Image Analysis, 1 (1997), pp. 195–206.

[16] J. Morel and S. Solemini, *Variational Methods in Image Segmentation*, Birkhäuser Boston, Cambridge, MA, 1995.

[17] P. Perona and J. Malik, *Scale-space and edge detection using anisotropic diffusion*, IEEE Trans. Pattern Anal. Mach. Intell., 12 (1990), pp. 629–639.

[18] M. Piccioni, S. Scarlatti, and A. Trouvé, *A variational problem arising from speech recognition*, SIAM J. Appl. Math., 58 (1998), pp. 753–771.

[19] A. Roche, G. Malandain, X. Pennec, and N. Ayache, *The correlation ratio as a new similarity measure for multimodal image registration*, in Medical Image Computing and Computer-Assisted Intervention, W.M. Wells, A.C.F. Cochester, and S. Delp, eds., Lecture Notes in Comput. Sci. 1496, Springer, New York, 1998, pp. 1115–1124.

[20] L. Rudin, S. Osher, and E. Fatemi, *Nonlinear total variation-based noise removal algorithms*, Phys. D, 60 (1992), pp. 259–268.

[21] P. Thompson and A. Toga, *Anatomically driven strategies for high-dimensional brain image registration and pathology*, in Brain Warping, Academic Press, New York, 1998, pp. 311–336.

[22] P. Viola and W.M. Wells, *Alignment by maximization of mutual information*, Internat. J. Computer Vision, 24 (1997), pp. 137–154.

[23] W.M. Wells, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis, *Multimodal volume registration by maximization of mutual information*, Medical Image Analysis, 1 (1996), pp. 35–51.