SIAM J. SCI. COMPUT.
Vol. 23, No. 4, pp. 1077–1093
© 2001 Society for Industrial and Applied Mathematics

ITERATIVE MULTIGRID REGULARIZATION TECHNIQUES FOR IMAGE MATCHING*

STEFAN HENN† AND KRISTIAN WITSCH‡

Abstract. In this paper, we consider the problem of matching images, i.e., to find a deformation u, which transforms a digital image into another such that the images have nearly equal gray values in every image element. The difference of the two images is measured by their L_2-difference, which should be minimized. This yields a nonlinear ill conditioned inverse problem for u, so the numerical solution is quite difficult. A Tikhonov regularization method is considered to rule out discontinuous and irregular solutions to the minimization problem. An important problem is a proper choice of the regularization parameter α. For the practical choice of α, we use iterative regularization methods based on multigrid techniques. To obtain a suitable initial guess, we use an approach similar to the full multigrid (FMG) developed by Brandt [*Math. Comp.,* 31 (1977), pp. 333–390]. The algorithms have optimal complexity: the amount of work is proportional to the number of picture elements. Finally, we present some experimental results for synthetic and real images.

Key words. image processing, multigrid methods, nonlinear ill-posed problems, Tikhonov regularization

AMS subject classifications. 65J20, 65N55, 68U10

PII. S106482750037161X

1. Introduction.

An important problem in two- and three-dimensional medical image analysis is to match two similar images resulting from different imaging modalities. In brain research, which was our starting point, it is necessary for the analysis of the organization and variation in the structure of human brains. A good survey of a part of the practical applications is given in [6]. In this paper, we consider a gray-level based mapping method of two similar images. The description of the problem and the numerical algorithms for its solution will be given for the two-dimensional problem. The extension to three dimensions is obvious and has been done.

In two dimensions, the matching leads to the following problem: find a displacement vector $u = (u_1, u_2)^t$ (whose components are functions of the variables $x = (x_1, x_2)^t$) that transforms one image T, called the template, into another image R, called the reference, so that

$$R(x) = T(x - u(x)). \tag{1.1}$$

In continuous two-dimensional variables, the reference R and the template T can be represented by functions $T, R : \Omega \subset \mathbb{R}^2 \longrightarrow \mathbb{R}^{\geq 0}$ which associate with the pixel (picture element) $(x_1, x_2) \in \mathbb{R}^2$ its intensity (gray level), $T(x_1, x_2)$, and $R(x_1, x_2)$, respectively. In our example, Ω will simply be the unit square $[0, 1]^2 \subset \mathbb{R}^2$. A digital image B^h is a finite-dimensional approximation of a continuous image B with $B^h : \Omega_h \to [0, g_{max}] \subset \mathbb{N}^{\geq 0}$ and maximal gray-level g_{max}.

We define a map $F = T \circ \phi$, with $\phi(x) = x - u(x)$, from the space of displacements

*Received by the editors May 2, 2000; accepted for publication (in revised form) June 5, 2001; published electronically October 11, 2001.

http://www.siam.org/journals/sisc/23-4/37161.html

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

‡Institut für Angewandte Mathematik, Heinrich-Heine Universität Düsseldorf, Universitätsstraße 1, D-40225 Düsseldorf, Germany (witsch@am.uni-duesseldorf.de).

$\mathcal{X}$ into the set $\mathcal{Y}$ of images

$$F : \mathcal{D}(F) \subset \mathcal{X} \longrightarrow \mathcal{Y}, \qquad F : u \longmapsto T(x - u(x)). \tag{1.2}$$

We consider the subset

$$\mathcal{D}(F) = \left\{ u \in \mathcal{X} \;\middle|\; x - u(x) \in \Omega \quad \forall x \in \Omega \right\}$$

of $\mathcal{X}$. This means that T, and therefore F, are defined. (This restriction can be avoided by extending T to a larger domain.) Thus, the gray-value matching problem (1.1) can be identified with an inverse problem

$$\text{find } u \text{ in (1.2), so that (1.1) holds approximately.}$$

In general, this problem is ill-posed in the sense of Hadamard. The information provided from R and the model (1.1) are not sufficient to ensure existence, uniqueness, and stability of a solution u.

In practical applications, only a noisy version R^δ of the exact data R is given with

$$||R - R^\delta|| \leq \delta.$$

In our application, δ is unknown, and we make no assumptions about statistics of the noise. In the case of perturbed data R^δ, the computation of a proper approximation u^δ of (1.1) requires regularization methods.

In section 2, we will present more precisely the minimization problem studied here, as well as the assumptions to impose on the model. We use a Tikhonov regularization for the so-called output least squares functional. It uses a regularization term which favors smooth transformations. This leads to a nonlinear variational problem similar to many other approaches in digital image processing or pattern recognition; cf. [1], [3], [7], [11], [12], [13], and [15].

The first attempts to deal with the issue of medical image matching can be traced back to Bajcsy and Kovacic [5]. These ideas were further developed in [8] using successive overrelaxation for a linearized model and in [1] using Fourier and wavelet representations of a nonlinear functional. In [11] a Landweber iteration scheme was developed, including the multigrid correction scheme (CS) and a trust region method for one-dimensional minimization.

The minima of the Tikhonov functional depend significantly on the choice of the regularization parameter α. Theoretically, decreasing α should give a better agreement of the images, but in practice a too small α leads to strong artifacts because of the influence of high-frequency structures in the image data. Increasing α removes these artifacts and allows only smoother transformations. For further increasing α, the matching of the images becomes worse.

Our aim is to develop methods for the efficient and robust solution of the matching problem and to find a good value of the regularization problem without a priori information.

These methods are described in section 4. In a first approach, a standard multigrid method and the choice of a suitable regularization parameter are combined. The second method follows the solution curve for decreasing α. One starts with a relative large α, which is helpful for the solution method. Then one computes for decreasing α minimal solutions of the functional consisting of the output least squares part and

a regularizing part applied to the difference from the previous solution. This process is stopped at the point where the output least squares functional increases.

Another important idea is to use a nested iteration or full multigrid (FMG) approach as introduced by Brandt [4]. The solution process starts with the same problem but a coarser resolution. It is easier to solve the nonlinear problem there and to find global minima. This solution is the starting point for the next finer resolution and usually needs only relatively small corrections.

The Euler–Lagrange equations of the Tikhonov functional are a coupled system of nonlinear PDEs which depend on the regularization parameter. Its discretization and approximation is described in section 3. Finally, we present some experimental results for synthetic and real images in section 6.

2. Tikhonov regularization of the nonlinear inverse problem. To find an approximate solution of (1.1), we measure the $L_2(\Omega)$ difference of the two images $h(u) = T(x-u) - R(x) = F(u) - R(x)$ by

$$D(u) = \langle h(u), h(u)\rangle_{L_2(\Omega)} = ||h(u)||^2_{L_2(\Omega)} = \int_\Omega (F(u(x)) - R(x))^2 d\Omega.$$

Here, $D(u)$ is the so-called output least squares functional with given data $R(x)$. The inverse problem will be solved by approximatively minimizing $D(u)$, which is a standard method for inverse problems. We can add further terms related, for example, to edges without essential changes in the algorithm presented here.

2.1. Tikhonov regularization. To rule out discontinuous and irregular solutions to the minimization problem, it is necessary to introduce a smoothing or regularizing term $\alpha||u||^2_{\mathcal{X}}$ with a parameter $\alpha > 0$ for the deformation u. This means that one minimizes the Tikhonov functional

$$J_\alpha(u) = ||h(u)||^2_{L_2(\Omega)} + \alpha||u||^2_{\mathcal{X}} \quad \text{for} \quad u \in \mathcal{D}(F) \tag{2.1}$$

with an appropriate norm $||\cdot||_{\mathcal{X}}$ for the space $\mathcal{X}$.

The statistical model proposed by Amit, Grenander, and Piccioni [2] for image restoration leads us to choose $||u||^2_{\mathcal{X}} = a(u,u)$ with a bilinear form a and to search for a solution among the minima of the energy

$$J_\alpha(u) = ||h(u)||^2_{L_2(\Omega)} + \alpha a(u,u) \quad \text{over} \quad \mathcal{D}(F). \tag{2.2}$$

The parameter $\alpha \in \mathbb{R}^+$ allows us to balance the influence of both terms in the Tikhonov functional $J_\alpha(u)$. For the brain data set, we use a regularizing term of the form

$$a(u,v) = \int_\Omega \left[\sum_{i,j=1}^{2} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} \right) \right] d\Omega \tag{2.3}$$

defined on $\mathcal{X} \subset H^1(\Omega) \times H^1(\Omega)$. This model assumes that the difference between the images comes from an elastic deformation of the brain. The elastic constants are chosen so that the changes in volume are maximal, which corresponds to longitudinal stretch without lateral shrink. The bilinear form 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 boundary conditions introduced in section 2.2.

This is an appropriate model for the case that the template image T is obtained by slicing and scanning the brain as described in [5]. Other bilinear forms on $H^1(\Omega) \times H^1(\Omega)$ as the Laplacian would give similar results but have no coupling between the directions as (2.3). Regularization in $H^2(\Omega) \times H^2(\Omega)$ as by the biharmonic operator enforces too much smoothness of the displacements.

Total variation based regularization as in [7], [13] is not appropriate in this case, but useful for other applications, and will be investigated in the future.

2.2. Minimization by a variational equation. With the described form of regularization in the functional J_α and $f(u) = -\frac{1}{\alpha} h'(u) \cdot h(u)$, the Euler–Lagrange equation for the minimizer u is given by the nonlinear variational equation

$$a(u, \phi) = \langle f(u), \phi \rangle_{L_2(\Omega)} \quad \forall \phi \in (H^1(\Omega))^2 = H^1(\Omega) \times H^1(\Omega). \tag{2.4}$$

For α very small, this is a singular perturbed problem which is difficult to solve.

Equation (2.4) is the weak form of a nonlinear coupled PDE. For $\mathcal{X} = (H^1(\Omega))^2$, one obtains Neuman boundary conditions. In our application, it is appropriate to use $\mathcal{X} = (H_0^1(\Omega))^2$ which yields Dirichlet boundary conditions

$$u(x) = 0 \quad \text{for} \quad x \in \partial\Omega.$$

(In principle, one has to enforce $u \in \mathcal{D}(F)$, but this was no problem in our application.) The minimization of the Tikhonov functional is therefore equivalent to finding a solution of the boundary value problem

$$E(u(x)) = \begin{cases} -\alpha(\Delta u(x) + \operatorname{grad}(\operatorname{div} u(x))) - f(u(x)) &= 0 \quad \text{for } x \in \Omega, \\ u(x) &= 0 \quad \text{for } x \in \partial\Omega. \end{cases} \tag{2.5}$$

3. Discretization, approximation, and solution methods of the nonlinear PDE. For the purpose of the numerical solution of (2.5), we approximate the infinite-dimensional space $\mathcal{X}$ by a finite-dimensional space $\mathcal{X}_n$.

3.1. Discretization. The image T^h obtained from the continuous image T can be considered as a function which is piecewise constant on each pixel of a regular square grid G_h. For the discretization of the PDE (2.5), we use G_h with the pixel-centered gridpoints as in Figure 1:

$$G_h = \{x \in \mathbb{R}^2 : x = (x_i, x_j) = (h/2 + ih, h/2 + jh),\ i, j = 0, 1, \ldots, n-1\}.$$

The displacement vector is a grid function $u_h(x_h) = (u_{1,h}, u_{2,h})^t \in (\mathcal{F}(G_h))^2 = \mathcal{F}(G_h) \times \mathcal{F}(G_h)$ on G_h. (In three-dimensional magnetic resonance (MR), there is a sampling difference between the $x-y$-plane and the z-direction. In practice, it is not necessary to take this into account for the adaptation process. It could be done, but then the multigrid (MG) components would have to be modified.)

3.2. Computation of the nonlinear functional $f(u(x))$. The nonlinear functional $f(u(x))$ in (2.5) has to be evaluated by discretization. With $\phi(u) = \phi(x, u) = x - u(x)$, one gets

$$f(u) = h'(u) \cdot h(u) = J_{T^h}(\phi(u)) \cdot J_\phi(u) \cdot h(u) = -\nabla T^h(\phi(u)) \cdot h(u),$$

since $J_\phi(u) = -I_2$, the 2×2 identity, and $J_{T^h}(x) = T^{h\prime}(x) = \nabla T^h(x)^t$. For determining a second order approximation of ∇T^h, the right-hand side $f_h(u_h) = (f_{1,h}, f_{2,h})^t \in$

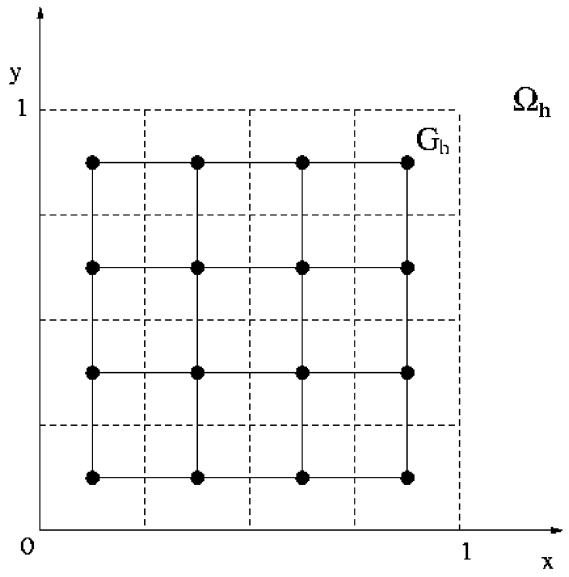

FIG. 1. *The pixel-centered grid G_h in two-dimensions for 4×4 pixels.*

$\mathcal{F}(G_h) \times \mathcal{F}(G_h)$ for an inner point (x_i, x_j) of G_h is computed with central differences by

$$f_h(u_h(x_i,x_j)) = \frac{1}{2}\left(\frac{T^h_{i+1,j} - T^h_{i-1,j}}{h}, \frac{T^h_{i,j+1} - T^h_{i,j-1}}{h}\right)^t (T^h_{i,j} - R^h_{i,j}), \tag{3.1}$$

where t denotes the transpose, $T^h_{i,j} = T^h(x_i - u_{1,h}(x_i,x_j), x_j - u_{2,h}(x_i,x_j))$ the deformed template image, and $R^h_{i,j} = R^h(x_i,x_j)$ the reference image.

3.3. Approximation of the PDE. The PDE (2.5) is discretized by the finite difference method using second order approximations. This is done by replacing the partial derivatives by corresponding difference quotients. This yields the following second order stencils:

$$-\Delta u \;=\; -u_{xx} - u_{yy} = \frac{1}{h^2}\begin{bmatrix} 0 & -1 & 0 \\ -1 & +4 & -1 \\ 0 & -1 & 0 \end{bmatrix} u_h + \mathcal{O}(h^2)$$

and

$$-u_{xy} = \frac{1}{4h^2}\begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix} u_h + \mathcal{O}(h^2).$$

Therefore, the discrete operator $L_h = -\Delta_h - \mathrm{grad}_h(\mathrm{div}_h)$ is represented by

$$L_h(u_h) = \begin{cases} \dfrac{1}{h^2}\begin{bmatrix} 0 & -1 & 0 \\ -2 & 6 & -2 \\ 0 & -1 & 0 \end{bmatrix} u_{1,h} \;+\; \dfrac{1}{4h^2}\begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix} u_{2,h}, \\ \dfrac{1}{h^2}\begin{bmatrix} 0 & -2 & 0 \\ -1 & 6 & -1 \\ 0 & -2 & 0 \end{bmatrix} u_{2,h} \;+\; \dfrac{1}{4h^2}\begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix} u_{1,h} \end{cases} \tag{3.2}$$

for $u_h = (u_{1,h}, u_{2,h})^t \in (\mathcal{F}(G_h))^2$. By using (3.2) for all inner gridpoints $x_h \in G_h$ and $u_h = 0$ for all boundary gridpoints, together with the discretized form of the nonlinear functional $f(u)$ as described in (3.1), one obtains a nonlinear system for the unknown values of the translation vector

$$E_h(u_h(x)) = \alpha L_h(u_h(x)) - f_h(u_h(x)) = 0. \tag{3.3}$$

This is the discretized form of the boundary value problem (2.5). The nonlinear system in two dimensions for typical images with 256×256 pixels has 2^{17} equations and unknowns. The corresponding system in three dimensions with $256 \times 256 \times 128$ voxel has 3×2^{23} equations and unknowns. This makes clear that it is extremely important to use very efficient solvers. Otherwise, in particular the three-dimensional problem cannot be solved.

3.4. Coarse grid problems and intergrid transfers. In the discretized PDE (3.3), the nonlinear functional $f_h(u_h)$ is naturally defined on the fine grid G_h by the pixel size of the digital images T^h and R^h (section 3.2). To use efficient solution methods like the multigrid full approximation scheme (FAS) [4], [14], for the nonlinear PDE we have to define the functional $f_h(u_h)$ also on coarse grids.

The basic idea of the FAS is to smooth the errors of the solution such that they can be approximated on a coarser grid. On the coarse grid, a defect equation is solved, and then the coarse grid corrections are interpolated back to the fine grid, where the errors are smoothed again. The components of the FAS for this problem are described in the following.

By coarsening the digital images, a corresponding nonlinear functional $f_H(u_H)$ can be defined on coarse grids $G_{2^l h}$ with $l > 0$. This means that we approximate the infinite-dimensional spaces $\mathcal{X}$ and $\mathcal{Y}$ by a sequence of finite-dimensional subspaces $\{\mathcal{Y}_l\}_{l=0,1,2,\ldots}$ with $\mathcal{Y}_l = \mathcal{Y}_{2^l h}$ for the images and $\{\mathcal{X}_l\}_{l=0,1,2,\ldots}$ with $\mathcal{X}_l = \mathcal{X}_{2^l h}$ for the displacements. The sequences have the properties

$$\mathcal{X}_L \subset \mathcal{X}_{L-1} \subset \cdots \subset \mathcal{X}_0 \subset \cdots \subset \mathcal{X} \quad \text{and} \quad \mathcal{Y}_L \subset \mathcal{Y}_{L-1} \subset \cdots \subset \mathcal{Y}_0 \subset \cdots \subset \mathcal{Y}.$$

The images are coarsened by collecting several picture elements into one coarse picture element. This can be described by the operator $R^l_{l-1} : \mathcal{Y}_{l-1} \to \mathcal{Y}_l$. This corresponds to a lowpass filter in digital imaging. This procedure yields a sequence of coarser and coarser images with containing only information about corresponding coarse structure. On all grids, the discretization of the PDE is done in the same way as on the finest grid as described in section 3.1; cf. Figure 2.

Standard components for the intergrid transfers (injection, half weighting, and full weighting for restriction and linear and bilinear interpolation) in the FAS were implemented.

4. Approximate solution methods. We now introduce two different methods for obtaining approximate solutions to the nonlinear problem. We consider the finite-dimensional Tikhonov functional $J_{\alpha,h} : G_h \times G_h \to \mathbb{R}^{\geq 0}$. The functional is positive and continuous. Therefore the Tikhonov functional has at least one global minimum on $\mathcal{X}$ because $J_{\alpha,h} \to \infty$ for $||u||_{\mathcal{X}} \to \infty$.

In general, the minimum is not unique. Moreover, a solution of $J_{\alpha,h}$ depends crucially on the regularization parameter α. The aim of our methods is to find a suitable solution of the minimization problem, hopefully a global one, without any a priori information about the image data and about a suitable choice of α in (2.2).

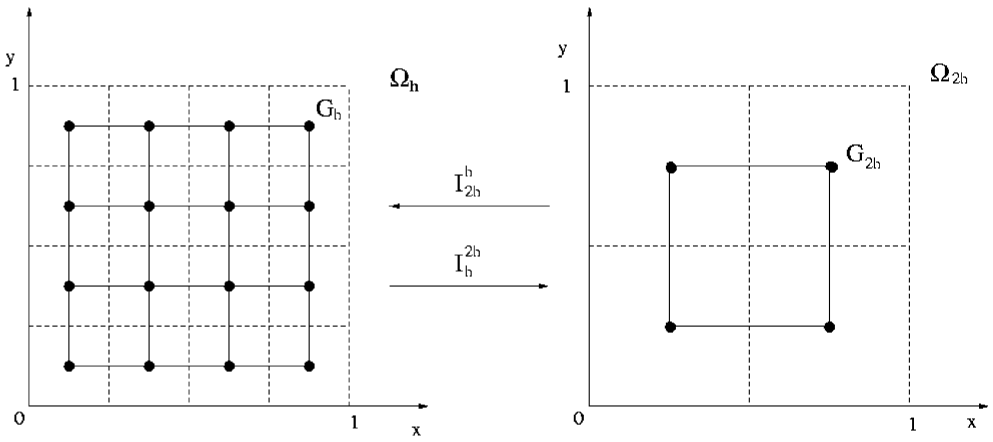

FIG. 2. *Intergrid transfers.*

The regularization parameter α determines the ratio between the influence of the image least squares differences $D(u)$ and the smoothness $a(u,u)$ of the displacements. For α too large, the influence of $D(u)$ is damped, and the matching is quite bad since the images behave as stiff materials. For α too small, the displacements are not smooth and are useless. Then, the numerical solution is difficult or even impossible. In particular, for $\alpha = 0$, each pixel could be mapped onto any other pixel with the same gray level.

In the first algorithm, we try to choose a deformation which is relatively independent of α. In the second algorithm, we solve a sequence of subproblems with decreasing parameter α. For each subproblem, we determine the parameter α within the iteration.

4.1. Algorithm 1: Modified FAS for minimizing output least squares functional. In this section, we will describe a method which incorporates the solution process and the choice of a suitable regularization in a modified FAS algorithm. We combine the multigrid idea and the minimization of the functional as follows.

The main components of every multigrid process are the relaxation and the coarse grid correction. As in standard multigrid methods, the coarse grid correction is determined by restricting the defect and the current approximation to the coarse grid. The coarse grid problem is solved approximately, the solution is interpolated to the fine grid and the solution is added to the fine grid approximation. However, in our case a proper scaling of the coarse grid correction and of the result of the relaxation yields significantly faster convergence, in particular for smaller α. Therefore, we proceed as follows.

On the fine grid an intermediate approximation $\overline{u}_h^{(k+1)}$ is computed by a nonlinear relaxation method for (2.5). We use a nonlinear Jacobi relaxation. With respect to the discretization and approximation of (2.5), we get

$$\frac{6}{h^2}u_{1,h}^{(n+1)} + f_{1,h}(u^{(n+1)}) = \frac{1}{h^2}\begin{bmatrix} & 1 & \\ 2 & & 2 \\ & 1 & \end{bmatrix} u_{1,h}^{(n)} + \frac{1}{4h^2}\begin{bmatrix} -1 & & 1 \\ & & \\ 1 & & -1 \end{bmatrix} u_{2,h}^{(n)}$$

$$\frac{6}{h^2}u_{2,h}^{(n+1)} + f_{2,h}(u^{(n+1)}) = \frac{1}{h^2}\begin{bmatrix} & 2 & \\ 1 & & 1 \\ & 2 & \end{bmatrix} u_{2,h}^{(n)} + \frac{1}{4h^2}\begin{bmatrix} -1 & & 1 \\ & & \\ 1 & & -1 \end{bmatrix} u_{1,h}^{(n)}$$

with initial guess $u_h^{(0)} \in \mathcal{D}(F)$. (On the coarse grids, the restricted defect of the next finer grid has to be added to the right-hand side.)

Due to the dependence on α, this approximation may result in a bad matching, as described above. To correct this, the approximation is modified by

$$u_h^{(k+1)} = u_h^{(k)} + \omega \cdot \left(\overline{u}_h^{(k+1)} - u_h^{(k)} \right).$$

The parameter ω is the solution of the one-dimensional minimization problem

$$\text{find } \omega \in \mathbb{R} \quad \text{so that} \quad \omega = \arg \min_{\overline{\omega} \in \mathbb{R}^{>0}} D\left(u_h^{(k)} + \overline{\omega} \cdot \left(\overline{u}_h^{(k+1)} - u_h^{(k)} \right) \right).$$

This approximation and its defect

$$d_h = rhs_h - E_h\left(\overline{u}_h^{(k+1)} \right)$$

are transferred to the coarse grid as in standard multigrid. On the fine grid, the right-hand side is given by $rhs_h = 0$ and as described in (4.2) on the coarse grid. On the coarse grid, we first choose a nonlinear relaxation scheme

$$\overline{u}_H = u_H + v_H = \mathcal{R}_H(u_H, v_H, d_H) \tag{4.1}$$

for the resulting nonlinear discrete system

$$E_H(u_H + v_H) = d_H + E_H(u_H) \tag{4.2}$$

with

$$u_H = \left(u_1^{(1,1)}, \ldots, u_1^{(n-1,n-1)}, u_2^{(1,1)}, \ldots, u_2^{(n-1,n-1)} \right) \in \mathcal{F}(G_H) \times \mathcal{F}(G_H)$$

and defect $d_H \in \mathcal{F}(G_H) \times \mathcal{F}(G_H)$ on the coarse grid to compute the coarse grid correction v_H in the full approximation scheme.

Again, we choose a parameter ω as the solution of the one-dimensional minimization-problem

$$\text{find } \omega \in \mathbb{R} \quad \text{so that} \quad \omega = \arg \min_{\overline{\omega} \in \mathbb{R}^{>0}} D(u_H + \overline{\omega} \cdot v_H). \tag{4.3}$$

Then the coarse grid correction is

$$\overline{v}_H = \omega \cdot v_H$$

and

$$u_h^{new} = u_h + I(\overline{v}_H).$$

In standard multigrid, a similar scaling is known, but $J_{\alpha,h}$ would be minimized. Here, the regularization term is used in determining the direction v_H but omitted in choosing ω. This reduces the influence of α on the final solution.

This two-level algorithm can be converted into a multigrid algorithm in the same way as in a standard FAS by recursive solution of the coarse grid problem by the two grid method.

4.2. Algorithm 2: Iterative Tikhonov regularization. In the previous section, we have described an algorithm which computes an approximation for a given parameter α but which reduces the influence of α on the solution by a one-dimensional minimization problem within a FAS.

In this section, we present an algorithm which is more complicated than algorithm 1 and is in some respect similar to the solution with the L-curve criterion investigated by Hansen [9] and Hansen and O'Leary [10]. The L-curve criterion is a practical method for choosing the regularization parameter α for linear inverse problems. It was observed that the log-log-graphic presentation of the curve $L = \{x(\alpha), y(\alpha)\}$ with the norm $x(\alpha)$ of the regularized solution $u_\alpha(x)$ of (2.2) and the corresponding residual norm $y(\alpha)$ can be useful in studying the problem.

The so-called L-curve method derives its name from the plot of the graph $\{x(\alpha), y(\alpha)\}$, which resembles the shape of the letter "L." Here, the basic idea is that a point on the graph near the "corner" of the "L" represents a reasonable compromise between data fit and smallness of the regularizing term. In our situation, the use of the L-curve criterion is dissatisfying due to the following reasons:

- The description of the L-curve criterion is for linear problems. (This means that F in (1.2) is a linear operator.) The idea can be generalized to nonlinear problems, but in general we saw no useful "L."
- For the iterative solution and the singular perturbation problems, we cannot compute solutions for small α.
- An important role for minimizing an nonlinear functional is the choice of an initial guess. This is not taken into account within the L-curve criterion. In contrast to linear problems, there is no optimal regularization parameter for an optimal solution.
- The L-curve criterion seems to prefer too smooth solutions in our case with nonoptimal matching of the images; see Figure 3.

In the following, we introduce a method which may be viewed as an iterated Tikhonov regularization. The method consists of a sequence of minimization subproblems

$$\left\{ \min_{u \in \mathcal{D}(F)} \{||h(u)||^2_{L_2(\Omega)} + \alpha_k ||u - u^{(k)}||^2_{\mathcal{X}}\} \right\}_{k \in \mathbb{N}} \tag{4.4}$$

whose solutions $u = u_\alpha^{(k+1)}$ depend on the parameter $\alpha_k > 0$ and on the initial guess $u^{(k)}$. Each subproblem is well posed for α sufficiently large and can be solved efficiently by a FAS.

There are two ideas: The first is, not to use a very small α in order to simplify the solution process. The second is the use of $u - u^{(k)}$ in the regularization term diminishes the influence of the regularization since $u^{(k)} \approx u$ but allows only reasonable solutions. An important issue herein is the determination of the first iterate.

For the numerical implementation of the iterative Tikhonov regularization, we compute the solution $u_\alpha^{(1)}$ with relative large α and an initial guess $u^{(0)} \in \mathcal{D}(F)$; cf. section 5. This is done by solving the nonlinear Euler–Lagrange equation (2.5) with a FAS and initial guess $u^{(0)}$. For the next iteration, we reduce α by a factor $0 < \kappa < 1$ (e.g., $\kappa = 1/2$) and take the solution $u_\alpha^{(1)}$ as the initial guess.

We get a sequence of displacement fields $\{u^{(k)}\}_{k \in \mathbb{N}}$ which successively reduces the value of the nonlinear functional $D(u)$. We stop the iteration if $D(u^{(k+1)}) > D(u^{(k)})$ and get the following algorithm.

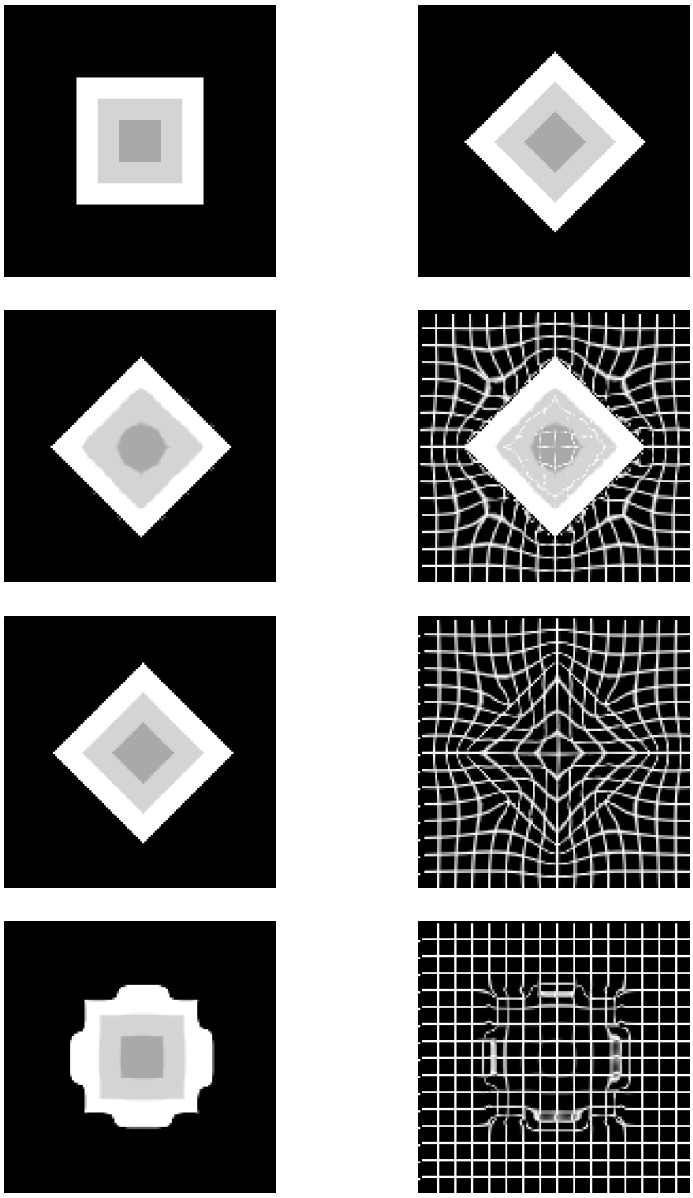

FIG. 3. *From left to right:* 1. *Template image* T. 2. *Reference image* R. 3. *Template deformed by the solution of algorithm* 1 *introduced in section* 4.1. 4. *Deformed template and superimposed deformed uniform grid for showing the deformation applied to the template.* 5. *Template deformed by the solution of algorithm* 2 *introduced in section* 4.2. 6. *Uniform grid deformed by the solution.* 7. *Template deformed by the solution of the L-curve criterion.* 8. *Uniform grid deformed by the solution of the L-curve criterion.*

Algorithm 2: Iterative Tikhonov regularization for minimizing $D(u)$.

$k = 0$; $u^{(k)} = u^*$; $\alpha_k = N \gg 0$;
repeat
  compute $u^{(k+1)} = \arg\min_{u \in \mathcal{D}(F)} ||h(u)||_2^2 + \alpha_k ||u^{(k)} - u^*||_{\mathcal{X}}^2$
  reduce $\alpha = \kappa \cdot \alpha$;
until $\left(D(u^{(k+1)}) > D(u^{(k)})\right)$.

5. Multiresolution minimization and multilevel image matching. Within the minimization of the Tikhonov functional (2.1) in sections 4.1 and 4.2, the subsets $\mathcal{X}_h \subset \mathcal{X}$ and $\mathcal{Y}_h \subset \mathcal{Y}$ are given by the number of pixels. This pixel size h is also the step size of the discretization of the PDE in section 3.

The use of a small discretization parameter h causes two difficulties:

- The adaption process mainly adapts structures with size h and neglects coarser structures. Therefore the minimization process finds local minima far from the global one.
- The minimization process requires many iterations with a fine resolution h.

Both difficulties can be avoided to a large extent by considering the Tikhonov functional also on coarser resolutions. The basic principle of this approach is that the image structures will be matched on that resolution on which they can be represented. This means that we approximate the infinite-dimensional spaces $\mathcal{X}$ and $\mathcal{Y}$ by a sequence of finite-dimensional subspaces, as in section 3.4. We get a sequence of minimization problems

$$(M_l) \qquad \min_{u_l \in \mathcal{X}_l} \left\{ \; ||h(u)||_2^2 + \alpha a(u,u) \; \right\}$$

defined on the spaces $\mathcal{X}_l$. With an interpolation operator $I_l^{l-1} : \mathcal{X}_{2^l h} \to \mathcal{X}_{2^{(l-1)} h}$, we transform the solution $u_l \in \mathcal{X}_l$ onto the next finer resolution. Then $I_l^{l-1}(u_l) \in \mathcal{X}_{l-1}$ is a suitable initial guess for the minimization of M_{l-1}.

Multilevel image matching.

$u_L = 0; l = L$;
repeat
  if $(l = L)$
    compute u_L^* with initial approximation u_L;
    $l = l - 1$;
  else
    $u_{l-1} = I_l^{l-1}(u_l^*)$;
    compute u_{l-1}^* with initial approximation u_{l-1};
  endif
  $l = l - 1$;
until l = 1.

6. Results. In order to conclude this paper, we present experimental results for synthetic as well as MR images. To demonstrate the principle and reliability of the iterative Tikhonov iteration scheme and the modified FAS described in sections 4.1 and 4.2, we consider the synthetic images shown in Figure 3. These images are quite different but can be adapted quite well, depending on the required smoothness.

Comparing the synthetic template image (leftmost in Figure 3) with the synthetic reference image (second left in Figure 3) shows a difference of $\frac{1}{N}||T(x) - R(x)||_{L_2} \approx 71$ with $N = 128^2$ pixels. After the iteration described in sections 4.1 and 4.2 on a

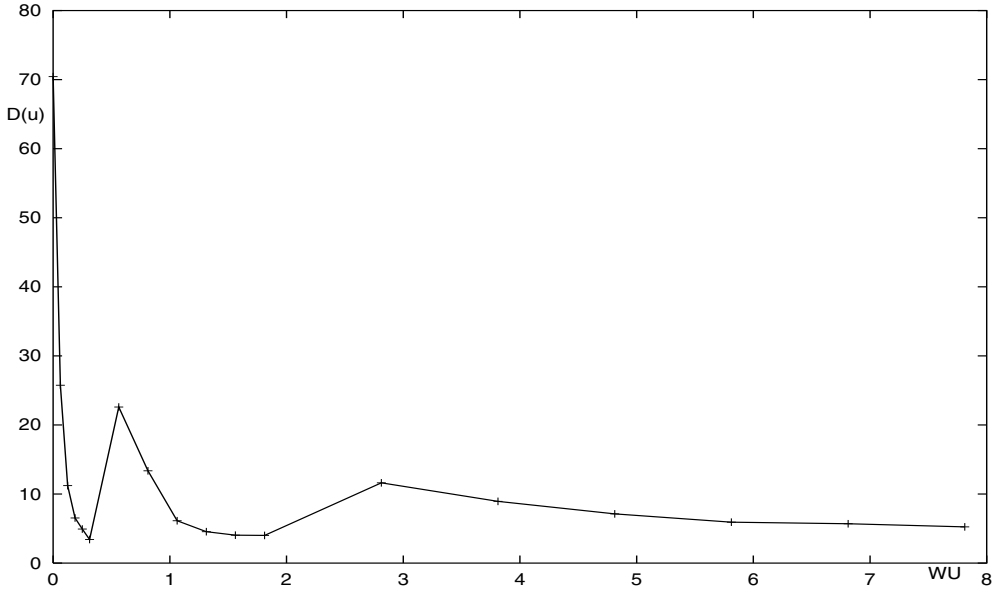

FIG. 4. *Least squares difference $D(u^{(k)})$ after each FAS-iteration (algorithm* 1 *in section* 4.1*) step for the example in Figure* 3.

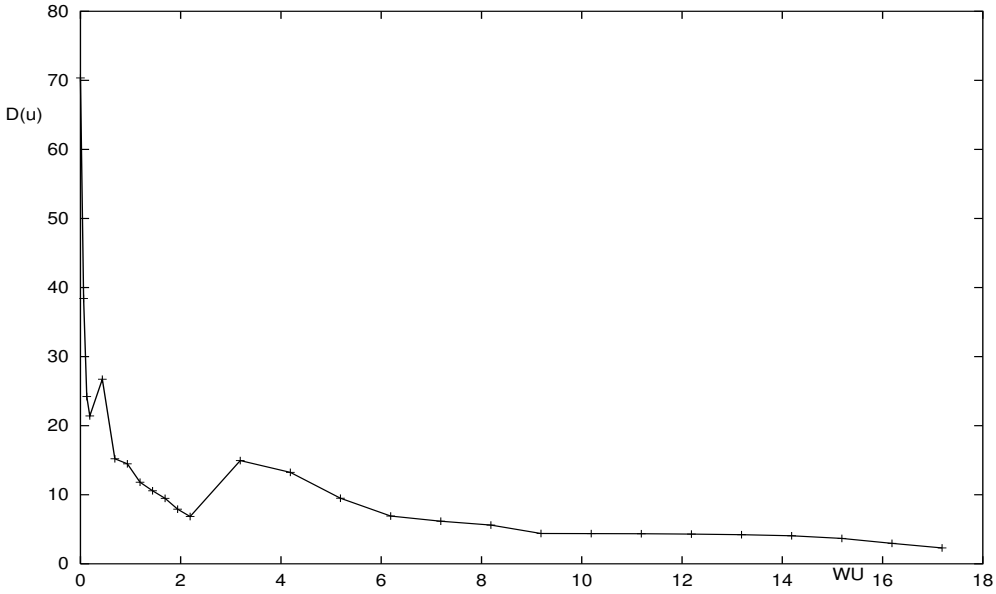

FIG. 5. *Least squares difference $D(u^{(k)})$ after each iteration step of the iterative Tikhonov regularization algorithm (algorithm* 2 *in section* 4.2*) for the example in Figure* 3.

sequence of the three subspaces $X_{1/32} \subset X_{1/64} \subset X_{1/128}$, we get a sequence of deformation fields $u^{(k)}$ on every subspace which reduces the difference $\frac{1}{N}||T(x-u^{(k)}(x))-R(x)||_{L_2}$ within the iteration. Figures 4 and 5 show for the algorithms the decreasing least squares difference on the three different resolutions applied to the example in

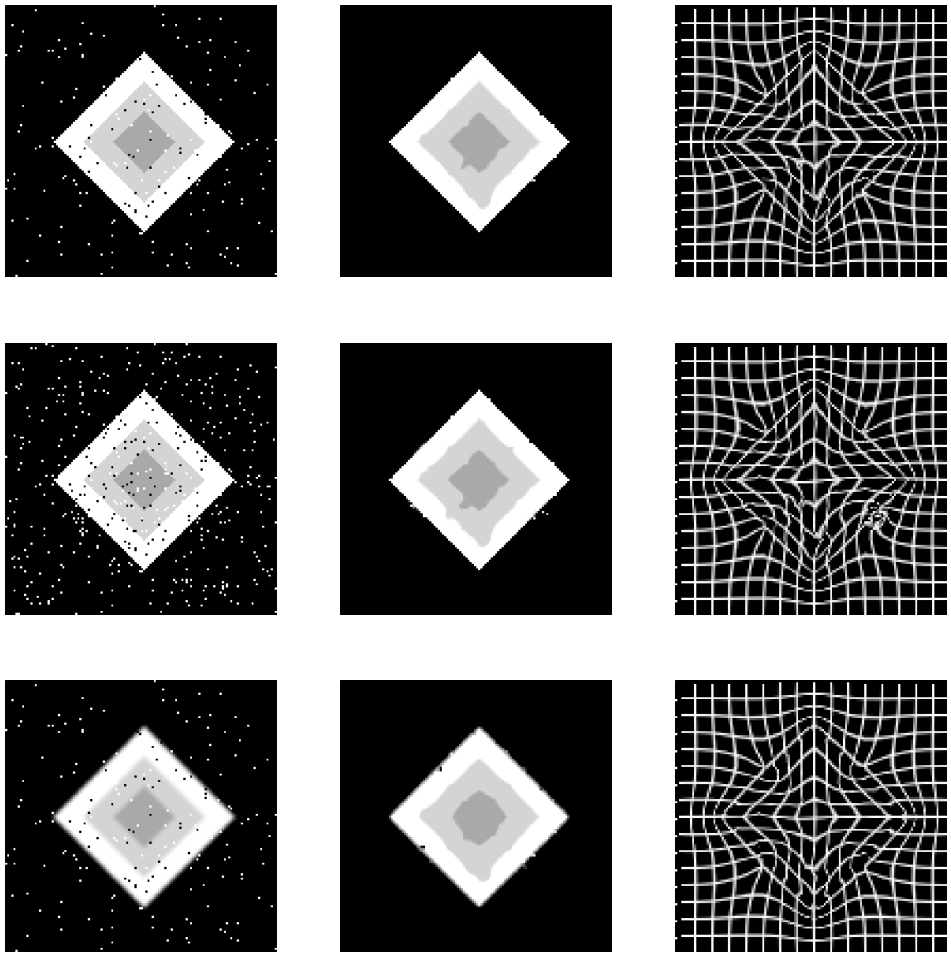

FIG. 6. *Images of the three experiments with different noise levels. Left: The reference with added noise R^δ (added noise from top to bottom:* 1. *2% salt and pepper noise* 2. *5% salt and pepper noise* 3. *blurred with 2% salt and pepper noise). Middle: Template image deformed by the solution of algorithm* 2. *Right: Uniform grid deformed by the solution of algorithm* 2.

Figure 3. The graphs are scaled by work units (WUs) on the finest grid, corresponding to the effort of a multigrid cycle.

In the examples, a WU of the first algorithm corresponds to approximately a $\frac{1}{3}$ WU of the second one. In all examples, we used a FAS with one relaxation before and one after the coarse grid correction for smoothing the error on three levels: full weighting restriction, bilinear interpolation, and the injection operator for the restriction of the image data and the approximate solutions.

In the next experiment, we analyze the dependence of the solution on the noise level $||R^\delta - R|| \leq \delta$ of the reference image R. For this, we added different types of noise to the synthetic reference images which appear in Figure 3. These images are shown in Figure 6. We observe that the results (in the middle of Figure 6) obtained by algorithm 2 are similar to the reference shown in Figure 3, except for artifacts caused by the added noise. The graphs of the decreasing defect functional $D(u)$ for these experiments scaled by WUs on the finest grid are shown in Figure 7. As expected, the

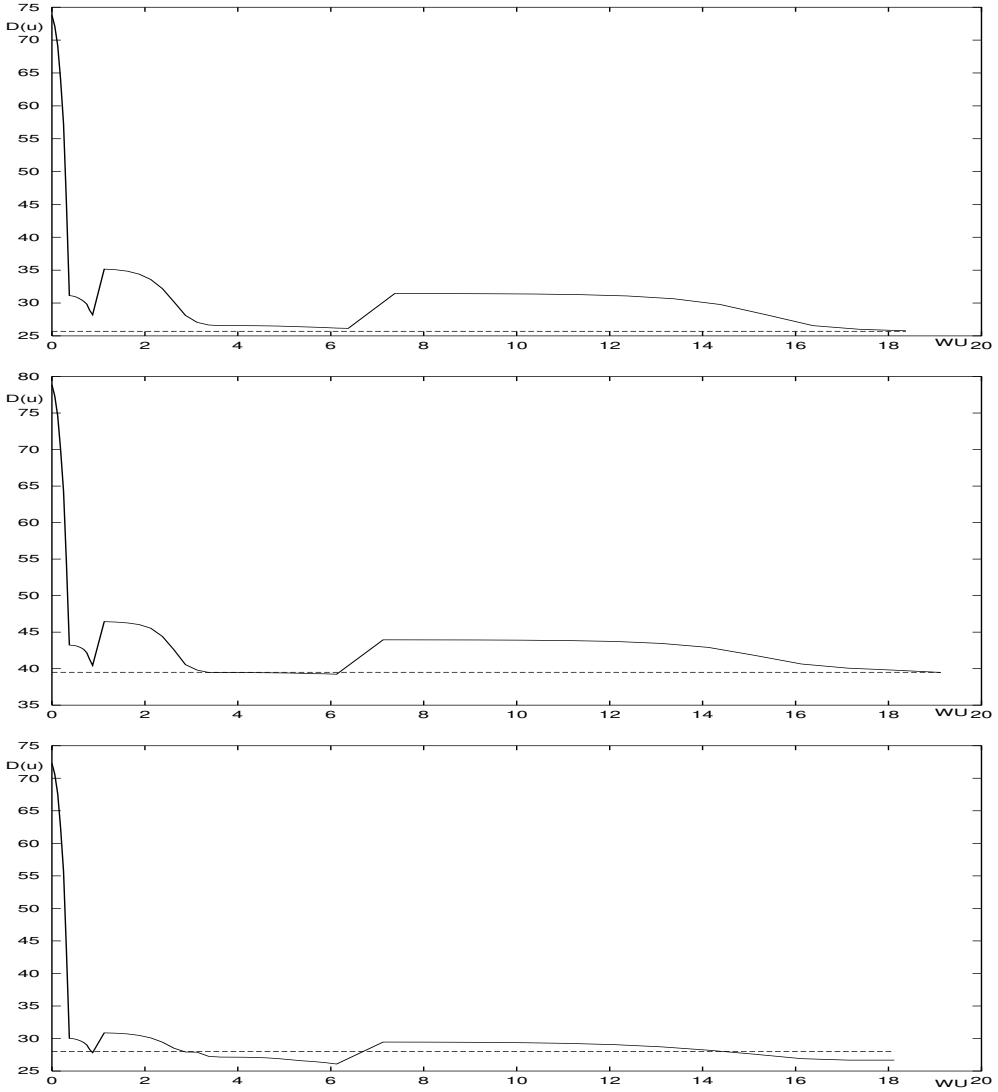

FIG. 7. *Least squares difference $D(u^{(k)})$ after each iteration step of algorithm* 2 *for the three experiments shown in Figure* 6 *and the noise level δ parallel to the abscissa.*

decreasing least squares functional on the three different resolutions can be reduced down to the noise level δ.

For estimation of the registration quality on MR slices, the iterative minimization algorithms for image matching are applied to a deformed MR template image $T(x)$ and a MR reference image $R(x)$, shown in Figures 8 and 9. Figure 10 shows the calculated result $T(x - u^*(x))$ deformed by the solution u^* of the minimization problem. The minimization problem is also solved on three subsets $X_{1/64} \subset X_{1/128} \subset X_{1/256}$. One can easily see in Figure 10 how the deformation field produces an image similar to the reference image. This can be stressed by the graphs in Figures 11 and 12. They show the decreasing least squares difference $D(u)$ between the images after each iteration step of the algorithms on different resolutions. The graphs are scaled by WUs on the finest grid.

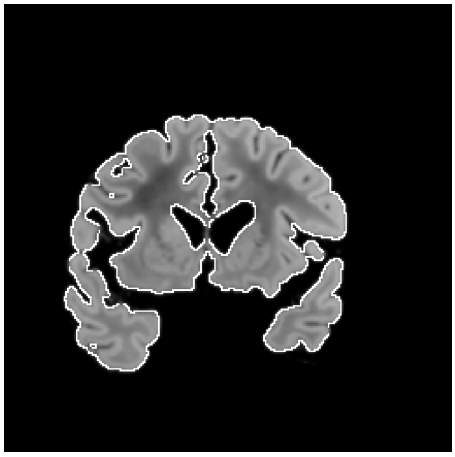

FIG. 8. *MR reference image with superimposed contour.*

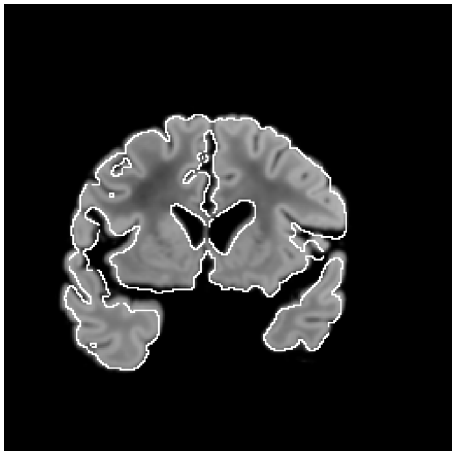

FIG. 9. *MR template slice with superimposed contour of the reference slice in Figure* 8.

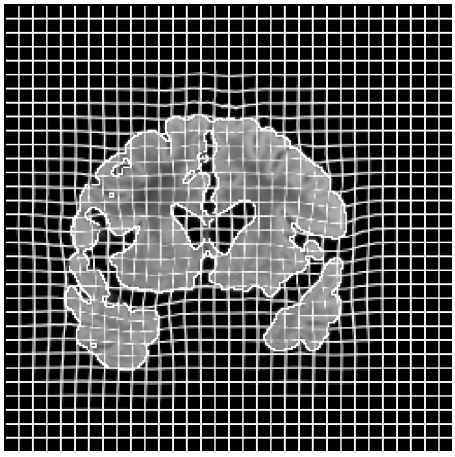

FIG. 10. *Deformed template with superimposed contour of the reference slice in Figure* 8 *and superimposed deformed uniform grid for showing the deformation applied to the template.*

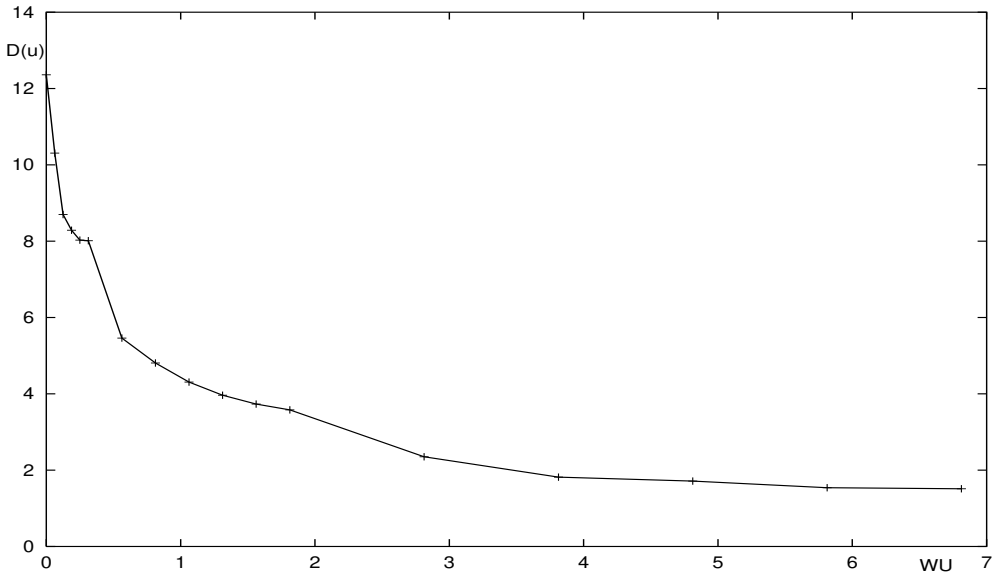

FIG. 11. *Least squares difference* $D(u^{(k)})$ *after each iteration step of algorithm* 1 *for the example in Figures* 8–10.

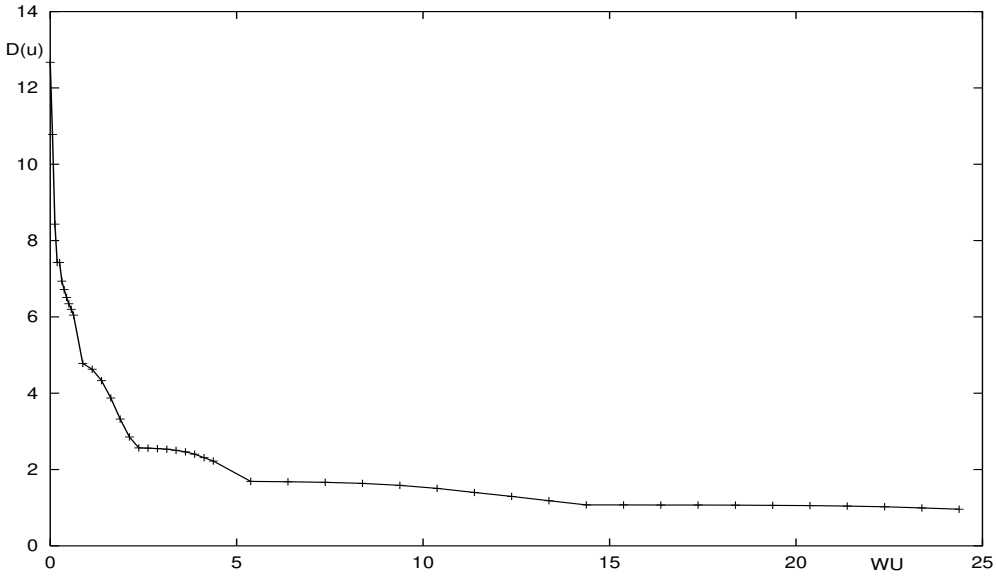

FIG. 12. *Least squares difference* $D(u^{(k)})$ *after each iteration step of algorithm* 2 *for the example in Figures* 8–10.

7. Conclusion. In this paper, we have introduced a nonlinear model for digital image matching. This inverse problem is ill-posed and its treatment requires careful use of regularizing techniques. For Tikhonov regularization, which aims at smooth solutions, a bilinear form is added. In section 4 we have presented two efficient methods for solving the problem.

The main advantages of these methods, compared to other known methods, are the self-controlled choices of the regularization parameter. In other papers [5], [1],

this problem is not discussed, and therefore in practical applications "trial an error" methods for the parameter choice are inevitable. On the other hand, the use of a FAS for solving the nonlinear Euler–Lagrange equations is an essential benefit of the presented methods. This results in computing time which is linear in the number of the picture elements. This optimal scaling is important for the treatment of large, in particular three-dimensional, problems which are in practice unsolvable only with relaxation methods. In order to compare our methods with other methods in literature, Bajcsy and Kovacic [5] use the Jacobi method to determine deformations of a template and Christensen et al. [8] use successive overrelaxation (SOR) with checkerboard update to compute viscous fluid deformations of a template. In practice, due to the h-dependence convergence rates of the Jacobi method and the SOR, these methods are rather slow. For instance, the method in [8] has an execution time of 7 days with a MIPS R440 processor and 10 hours using a massively parallel DECmpp 128×64 MasPar computer with a resolution of $128 \times 128 \times 100$ voxel. With this resolution, our algorithm takes (with a nonoptimized implementation in C++) approximately one hour on an Ultra-Sparc Workstation.

Acknowledgment. We would like to thank the referees for useful remarks on the first version of the manuscript.

REFERENCES

[1] Y. Amit, *A nonlinear variational problem for image matching*, SIAM J. Sci. Comput., 15 (1994), pp. 207–224.

[2] Y. Amit, U. Grenander, and M. Piccioni, *Structural image restoration thought deformable templates*, J. Amer. Statist. Assoc., 86 (1991), pp. 376–387.

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

[4] A. Brandt, *Multi-level adaptive solutions to boundary value problems*, Math. Comp., 31 (1977), pp. 333–390.

[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, vol. 30/no. 3, 1997.

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

[9] P.C. Hansen, *Analysis of discrete ill-posed problems by means of the L-curve*, SIAM Rev., 34 (1992), pp. 561–580.

[10] P.C. Hansen and D.P. O'Leary, *The use of the L-curve in the regularization of discrete ill-posed problems*, SIAM J. Sci. Comput., 14 (1993), pp. 1487–1503.

[11] S. Henn and K. Witsch, *A multigrid-approach for minimizing a nonlinear functional for digital image matching*, Computing, 64 (2000), pp. 339–348.

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

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

[14] K. Stüben and U. Trottenberg, *Multigrid Methods: Fundamental Algorithms, Model Problem Analysis and Applications*, Lecture Notes in Math. 960, Springer-Verlag, Berlin, 1982.

[15] D. Terzopoulos, *Image analysis using multigrid relaxation methods*, IEEE TPAMI, 2 (1986), pp. 129–139.