

# Multigrid solution of the nonlinear PDEs arising in elastic image registration with application to a group of monomodal images

Noppadol Chumchob<sup>a,\*</sup>, Sopida Jewprasert<sup>a</sup>, Chantana Chantrapornchai<sup>b</sup>

<sup>a</sup> Department of Mathematics, Faculty of Science, Silpakorn University, Nakorn Pathom 73000 Thailand

<sup>b</sup> Department of Computer Engineering, Faculty of Engineering, Kasetsart University, Bangkok 10900 Thailand

\*Corresponding author, e-mail: chumchob\_n@silpakorn.edu

Received 1 Aug 2015

Accepted 13 Dec 2016

**ABSTRACT:** This work addresses an important and yet outstanding issue for the numerical solution of the nonlinear partial differential equations arising in elastic image registration. We present an efficient multigrid method using local Fourier analysis and, in particular, its application to register a group of monomodal images. We demonstrate that the proposed multigrid method is fast, accurate, and reliable in delivering visually-pleasing registration results.

**KEYWORDS:** finite difference, image matching, local Fourier analysis

**MSC2010:** 65F10 65M55 68U10

## INTRODUCTION

In the last decades, there has been a dramatic growth in research and application of partial differential equation (PDE) based methods in image processing and computer vision. This is mainly due to the fact that these methods have remarkable advantages in both theory and computation, and that with the rapid progress of recent computer technology the solution of the resulting large systems of linear or nonlinear equations has become feasible using a normal PC.

PDE-based image registration methods have been actively and extensively studied and applied in the field of image analysis. The task of image registration is to find spatial correspondences between two given images, a so-called reference  $R : \Omega \subset \mathbb{R}^2 \rightarrow V \subset \mathbb{R}$  and a so-called template  $T : \Omega \subset \mathbb{R}^2 \rightarrow V \subset \mathbb{R}$ , where the image domain  $\Omega$  is assumed to be a rectangle. This is usually done by finding an optimal deformation  $\mathbf{u} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ,  $\mathbf{u} : \mathbf{x} \mapsto \mathbf{u}(\mathbf{x}) = (u_1(\mathbf{x}), u_2(\mathbf{x}))^T$ , such that the transformed template  $T_{\mathbf{u}}(\mathbf{x}) = T(\mathbf{x} + \mathbf{u}(\mathbf{x}))$  and the reference  $R$  are spatially matched, according to an image-to-image dissimilarity measure,  $\mathcal{D}(\mathbf{u}; T, R)$ . Without loss of generality, this work assumes that  $\Omega = [0, 1]^2 \subset \mathbb{R}^2$  and  $V = [0, 1]$  for 2D grey images.

If the image intensities of  $R$  and  $T$  are comparable (i.e., in a monomodal registration scenario), one may find various choices for  $\mathcal{D}$ . Probably the most popular choice for the dissimilarity measure is provided by the so-called sum of squared differences

$$\mathcal{D}(\mathbf{u}; T, R) = \frac{1}{2} \int_{\Omega} (T_{\mathbf{u}}(\mathbf{x}) - R(\mathbf{x}))^2 d\mathbf{x}.$$

The registration task is then to solve the minimization problem

$$\min_{\mathbf{u}} \{\mathcal{D}(\mathbf{u}; T, R)\}, \quad (1)$$

where  $\mathbf{u}$  is searched over a set of admissible functions  $\mathcal{U}$  minimizing  $\mathcal{D}$ . As is well known, the minimization of  $\mathcal{D}$  does not have a unique minimizer and it becomes necessary to impose a constraint on the solution  $\mathbf{u}$  via a regularizing term. In this work, the commonly used regularizing term of the form

$$\mathcal{E}^{\text{elas}} = \int_{\Omega} \frac{\mu}{4} \sum_{l,m=1}^2 \left( \frac{\partial u_m}{\partial x_l} + \frac{\partial u_l}{\partial x_m} \right)^2 + \frac{\lambda}{2} (\nabla \cdot \mathbf{u})^2 d\mathbf{x},$$

is used to ensure that the constructed deformation  $\mathbf{u}$  is unique and penalizes cracks, foldings, or other unwanted deformations. This term is known as the elastic regularizer resulting from the theory

of linear elasticity for homogeneous and isotropic materials, typically human tissue for the purpose of medical image registration. Here  $\mathbf{u}$  is assumed to be relatively small.  $\lambda \geq 0$  and  $\mu > 0$  are the so-called Lamé constants which reflect material properties. In contrast to other traditional regularizers, e.g., diffusion<sup>1</sup>, curvature<sup>2,3</sup>, or total variation based regularizers<sup>4</sup>, this regularizer couples the directions. The strength of the coupling depends on the choice of  $\lambda$ . By modifying (1), the desired deformation  $\mathbf{u}$  is a minimizer of the unconstrained optimization problem

$$\min_{\mathbf{u}} \{ \mathcal{J}_{\alpha}(\mathbf{u}; T, R) = \mathcal{D}(\mathbf{u}; T, R) + \alpha \mathcal{R}^{\text{elas}}(\mathbf{u}) \} \quad (2)$$

where  $\alpha > 0$  is a positive constant that compromises the quality of the similarity between  $T_{\mathbf{u}}$  and  $R$ , measured by  $\mathcal{D}$ , and the level of penalty for unwanted deformations, measured by  $\mathcal{R}$ . For simplicity, it is appropriate to use the search space  $\mathcal{U} = (H_0^1(\Omega))^2$  rather than  $\mathcal{U} = (H^1(\Omega))^2$ . Here the space  $H^1(\Omega)$  is the Sobolev space of functions in  $L^2(\Omega)$ .  $H_0^1(\Omega)$  is the subspace of  $H^1(\Omega)$  consisting of functions with zero boundary values.

According to the calculus of variations, the minimization of  $\mathcal{J}_{\alpha}$  is equivalent to finding a solution of the boundary value problem

$$\mathbf{f}(\mathbf{u}) - (\bar{\mu} \Delta \mathbf{u} + (\bar{\lambda} + \bar{\mu}) \nabla \text{div} \mathbf{u}) = \mathbf{0} \quad (3)$$

subject to the boundary conditions  $u_1(\mathbf{x}) = u_2(\mathbf{x}) = 0$  for  $\mathbf{x} \in \partial\Omega$ , where  $\mathbf{f}(\mathbf{u}) = (T_{\mathbf{u}} - R) \nabla_{\mathbf{u}} T_{\mathbf{u}}$  and  $\partial\Omega$  denotes the image boundary. Here  $\bar{\lambda} \equiv \alpha\lambda$  and  $\bar{\mu} \equiv \alpha\mu$ . Note that (3) is known as the Navier-Lamé equation.

Most commonly, the PDE in (3) is approximated by a finite difference scheme, which results in a large but sparse nonlinear algebraic system of equations for high-resolution digital images. For example, a registration problem of two  $4096 \times 4096$  mammography images leads to a discrete nonlinear system of equations in about  $32 \times 10^6$  unknowns. Thus reliable and computationally efficient solvers for such systems are essential in order to perform the registration in a reasonable amount of time. One of the most efficient techniques for solving such large systems is the so-called multigrid (MG) method. This approach uses a sequence of progressively coarser grids to accelerate the convergence of some basic relaxation process on the target grid. The basic relaxation method is referred to as the ‘smoother’, as its objective is to smooth the error in the current approximation relative to the computational grid. There are two general approaches

for solving discretized nonlinear partial differential equations (PDEs) by MG methods. One is to perform a global linearization, usually by Newton’s method, and solve the resulting linear system approximately by a linear MG algorithm; this is then repeated iteratively until the convergence of Newton’s method is obtained. The second approach, represented by the so-called full approximation scheme nonlinear MG (FAS-NMG) method, is to apply the method directly to the nonlinear system by performing only global and/or local linearization in the error-smoothing process. Convergence acceleration is then provided by nonlinear coarse-grid operators.

Over the years, variants of these two general categories have been proposed and developed in the field of PDE-based image registration methods. Although no proof exists, the latter approach has an efficiency similar to that of Newton’s method, but it is much more suitable due to two advantages. First, there is no requirement that the initial solutions are within the domain of attraction. Second, it is not necessary to compute and store the Jacobian in the latter approach, as is necessary in the Newton-based methods. As motivated by these advantages, the main aim of this paper is to propose an efficient MG solution for the boundary value problem given by (3). This work is closely related to that of Henn and Witsch<sup>5</sup> where an FAS-NMG solution is presented with a nonlinear Jacobi smoother. Although they share similar backgrounds (e.g., the same boundary value problem and MG methods), they are totally different by virtue of their smoothing processes and applications. In particular, the existence and efficiency of pointwise smoothing methods are investigated in this paper using local Fourier analysis (LFA) techniques.

The remaining parts of this paper are arranged as follows. The second section presents a finite difference method for (3) followed by the proposed MG solution with a potential smoother for solving the discrete nonlinear system in the third section. LFA techniques for the resulting discrete nonlinear operator and the proposed smoother also include in this section. The fourth section discusses implementation details in applying the proposed MG solution method for a group of monomodal images. The fifth section presents some experimental results to demonstrate the computational efficiency of the proposed MG solution in registering synthetic and real medical images as well as a group of real medical images. Concluding remarks are given in the last section.

### FINITE DIFFERENCE DISCRETIZATION

To start with, the PDE in (3) is rewritten as

$$\left. \begin{aligned} f_1(\mathbf{u}) - [(\bar{\lambda} + 2\bar{\mu})\partial_{x_1x_1}u_1 + \bar{\mu}\partial_{x_2x_2}u_1 \\ + (\bar{\lambda} + \bar{\mu})\partial_{x_1x_2}u_2] &= 0 \\ f_2(\mathbf{u}) - [(\bar{\lambda} + \bar{\mu})\partial_{x_1x_2}u_1 + \bar{\mu}\partial_{x_1x_1}u_2 \\ + (\bar{\lambda} + 2\bar{\mu})\partial_{x_2x_2}u_2] &= 0 \end{aligned} \right\} \quad (4)$$

where  $f_l(\mathbf{u}) = (T_u - R)\partial_{u_l} T_u$  for  $l = 1, 2$ . To discretize the PDEs represented in (4), let  $(u_l^h)_{i,j} = u_l^h(x_{1i}, x_{2j})$  be the grid function for  $l = 1, 2$  with the uniform grid spacing  $h = 1/n$ . The integer  $n = 1/h$  is the number of uniform intervals in both coordinate directions. Each grid point  $\mathbf{x}$  in the discretized domain  $\Omega^h$  is given by  $\mathbf{x} = (x_1, x_2)^T = (ih, jh)^T$  for  $0 \leq i, j \leq n$ . The partial derivatives are approximated by the standard second-order finite difference schemes. By the lexicographic column ordering of unknowns (starting from top-to-bottom and left-to-right), the discrete nonlinear system is given by

$$\mathbf{N}_h \mathbf{u}^h = \mathbf{G}_h[\mathbf{u}^h], \quad (5)$$

where  $(\cdot)_{i,j}$  is dropped for simplicity and

$$\mathbf{N}_h = \frac{1}{h^2} \begin{bmatrix} -\mathcal{L}_1^h & -\mathcal{P}^h \\ -\mathcal{P}^h & -\mathcal{L}_2^h \end{bmatrix}, \quad \mathbf{G}_h[\mathbf{u}^h] = \begin{bmatrix} -f_1^h(\mathbf{u}^h) \\ -f_2^h(\mathbf{u}^h) \end{bmatrix}.$$

The stencils corresponding to  $\mathcal{L}_1^h$ ,  $\mathcal{L}_2^h$ , and  $\mathcal{P}^h$  are given respectively by

$$\begin{aligned} \mathcal{L}_1^h &= \begin{bmatrix} 0 & 2\bar{\mu} + \bar{\lambda} & 0 \\ \bar{\mu} & -2(3\bar{\mu} + \bar{\lambda}) & \bar{\mu} \\ 0 & 2\bar{\mu} + \bar{\lambda} & 0 \end{bmatrix}, \\ \mathcal{L}_2^h &= \mathcal{L}_1^{h^T} = \begin{bmatrix} 0 & \bar{\mu} & 0 \\ 2\bar{\mu} + \bar{\lambda} & -2(3\bar{\mu} + \bar{\lambda}) & 2\bar{\mu} + \bar{\lambda} \\ 0 & \bar{\mu} & 0 \end{bmatrix}, \\ \mathcal{P}^h &= \frac{1}{4} \begin{bmatrix} -(\bar{\mu} + \bar{\lambda}) & 0 & \bar{\mu} + \bar{\lambda} \\ 0 & 0 & 0 \\ \bar{\mu} + \bar{\lambda} & 0 & -(\bar{\mu} + \bar{\lambda}) \end{bmatrix}. \end{aligned}$$

### THE PROPOSED MULTIGRID SOLUTION OF THE ASSOCIATED SYSTEM OF NONLINEAR EQUATIONS

The challenging part in applying the finite difference methods for nonlinear PDEs is how to effectively solve the associated system of nonlinear equations. Each system can be very large and is strongly coupled, but it has a few characteristics which are particularly amenable to MG methods. The main aim of this section is to propose an MG solution of the discrete nonlinear system as represented by (5).

### Local Fourier analysis and the measure of $h$ -ellipticity

Local Fourier analysis (LFA) is often applied to develop efficient MG methods. Although LFA was originally developed for linear elliptic PDEs with constant (or frozen) coefficients on infinite grids, it can also be applied to more general nonlinear PDEs with varying coefficients. LFA is based on the simplification that boundary conditions are neglected and all occurring discrete operators are extended to an infinite grid. On an infinite grid, the discrete nonlinear operator can be linearized (by freezing coefficients) and replaced locally by a new operator with constant coefficients<sup>6</sup>.

The  $h$ -ellipticity measure is an important quantity in the framework of LFA. It is often used to decide whether or not a certain discretization is appropriate for an MG treatment. A certain amount of  $h$ -ellipticity can be used as a sufficient condition for the existence of an efficient pointwise smoother. Because of the nonlinearity of the data terms  $f_l^h(\mathbf{u}^h) = f_l^h(u_1^h, u_2^h)$  in (5), the analysis will be performed locally on the linearized discrete system as given by

$$\mathbf{N}_h[\bar{\mathbf{u}}^h] \mathbf{u}^h = \mathbf{G}_h[\bar{\mathbf{u}}^h], \quad (6)$$

where  $\bar{\mathbf{u}}^h = (\bar{u}_1^h, \bar{u}_2^h)^T$  represents the current approximation obtained from solving (6), and

$$\begin{aligned} \mathbf{N}_h[\bar{\mathbf{u}}^h] &= \frac{1}{h^2} \begin{bmatrix} -\mathcal{L}_1^h + h^2\sigma_{11}[\bar{\mathbf{u}}^h] & -\mathcal{P}^h + h^2\sigma_{12}[\bar{\mathbf{u}}^h] \\ -\mathcal{P}^h + h^2\sigma_{21}[\bar{\mathbf{u}}^h] & -\mathcal{L}_2^h + h^2\sigma_{22}[\bar{\mathbf{u}}^h] \end{bmatrix} \end{aligned}$$

and

$$\mathbf{G}_h[\bar{\mathbf{u}}^h] = \begin{bmatrix} -f_1^h(\bar{\mathbf{u}}^h) + \sigma_{11}[\bar{\mathbf{u}}^h]\bar{u}_1^h + \sigma_{12}[\bar{\mathbf{u}}^h]\bar{u}_2^h \\ -f_2^h(\bar{\mathbf{u}}^h) + \sigma_{21}[\bar{\mathbf{u}}^h]\bar{u}_1^h + \sigma_{22}[\bar{\mathbf{u}}^h]\bar{u}_2^h \end{bmatrix}$$

are the discrete operators resulting from applying the first-order approximation to the data terms, i.e.,

$$f_l^h(\mathbf{u}^h) \approx f_l^h(\bar{\mathbf{u}}^h) + \sigma_{l1}(\bar{\mathbf{u}}^h)(u_1^h - \bar{u}_1^h) + \sigma_{l2}(\bar{\mathbf{u}}^h)(u_2^h - \bar{u}_2^h).$$

Here  $\sigma_{l1}(\bar{\mathbf{u}}^h) = \partial_{u_1} f_l^h(\bar{\mathbf{u}}^h) = (\partial_{u_1} T_{\bar{\mathbf{u}}^h}^h)(\partial_{u_1} T_{\bar{\mathbf{u}}^h}^h) + (T_{\bar{\mathbf{u}}^h}^h - R^h)(\partial_{u_1 u_1} T_{\bar{\mathbf{u}}^h}^h)$  and  $\sigma_{l2}(\bar{\mathbf{u}}^h) = \partial_{u_2} f_l^h(\bar{\mathbf{u}}^h) = (\partial_{u_1} T_{\bar{\mathbf{u}}^h}^h)(\partial_{u_2} T_{\bar{\mathbf{u}}^h}^h) + (T_{\bar{\mathbf{u}}^h}^h - R^h)(\partial_{u_2 u_1} T_{\bar{\mathbf{u}}^h}^h)$ . Because the image difference  $T_{\bar{\mathbf{u}}^h}^h - R^h$  is small for well-registered images ( $\bar{\mathbf{u}}^h$  is very close to the true solution),  $\sigma_{lm}(\bar{\mathbf{u}}^h)$  can be replaced by  $(\partial_{u_l} T_{\bar{\mathbf{u}}^h}^h)(\partial_{u_m} T_{\bar{\mathbf{u}}^h}^h)$  for  $m = 1, 2$ .

Let  $\Omega_\infty^h = \{\mathbf{x} \in \Omega \mid \mathbf{x} = (x_{1i}, x_{2j})^T = (ih, jh)^T, i, j \in \mathbb{Z}^2\}$  be the infinite grid and let

$\varphi_h(\theta, \mathbf{x}) = \exp(i\theta\mathbf{x}/h) \cdot \hat{\mathbf{I}}$  be grid functions, where  $\hat{\mathbf{I}} = (1, 1)^T$ ,  $\theta = (\theta_1, \theta_2)^T \in \Theta = (-\pi, \pi]^2$  and  $\mathbf{x} \in \Omega_h^\infty$ . It is important to remark that due to the local nature of LFA, the analysis should be applied to each grid point separately, i.e., it focuses on the local discrete system  $\mathbf{N}_h(\xi)\mathbf{u}^h(\xi) = \mathbf{G}_h(\xi)$  centred and defined only within a small neighbourhood of each grid point  $\xi = (x_i, y_j)$  and  $\mathbf{u}^h(\xi) = [u_1^h(\xi), u_2^h(\xi)]$ . Applying the discrete operator  $\mathbf{N}_h(\xi)$  to the grid functions  $\varphi_h(\theta, \mathbf{x})$ , i.e.,  $\mathbf{N}_h(\xi)\varphi_h(\theta, \mathbf{x}) = \hat{\mathbf{N}}_h(\xi, \theta)\varphi_h(\theta, \mathbf{x})$ , yields the Fourier component

$$\hat{\mathbf{N}}_h(\xi, \theta) = \frac{1}{h^2} \begin{bmatrix} -\mathcal{L}_1^h(\theta) + h^2\sigma_{11}(\xi) & -\mathcal{P}^h(\theta) + h^2\sigma_{12}(\xi) \\ -\mathcal{P}^h(\theta) + h^2\sigma_{21}(\xi) & -\mathcal{L}_2^h(\theta) + h^2\sigma_{22}(\xi) \end{bmatrix}. \quad (7)$$

Here  $\mathcal{L}_1^h(\theta) = -(2aw_1 + 2bw_2)$ ,  $\mathcal{L}_2^h(\theta) = -(2bw_1 + 2aw_2)$  and  $\mathcal{P}^h(\theta) = -c \sin \theta_1 \sin \theta_2$  denote, respectively, the Fourier components of the discrete operators  $\mathcal{L}_1^h$ ,  $\mathcal{L}_2^h$  and  $\mathcal{P}^h$ , where  $a = 2\bar{\mu} + \bar{\lambda}$ ,  $b = \bar{\mu}$ ,  $c = \bar{\mu} + \bar{\lambda}$ ,  $w_1 = 1 - \cos \theta_1$ , and  $w_2 = 1 - \cos \theta_2$ .

Following Ref. 6, the measure of the  $h$ -ellipticity is defined via  $\hat{\mathbf{N}}_h(\xi, \theta)$  as follows:

$$E_h(\mathbf{N}_h(\xi)) = \frac{\min\{|\det(\hat{\mathbf{N}}_h(\xi, \theta))| : \theta \in \Theta_{\text{high}}\}}{\max\{|\det(\hat{\mathbf{N}}_h(\xi, \theta))| : \theta \in \Theta\}}, \quad (8)$$

where  $\Theta_{\text{high}} = \Theta \setminus (-\pi/2, \pi/2)^2$  denotes the range of high frequencies and  $\det(\hat{\mathbf{N}}_h(\xi, \theta)) = [4ab((1 - \cos \theta_1)^2 + (1 - \cos \theta_2)^2) + 4(a^2 + b^2)(1 - \cos \theta_1)(1 - \cos \theta_2) - c^2 \sin^2 \theta_1 \sin^2 \theta_2]/h^4 + [c_1(1 - \cos \theta_1) + c_2(1 - \cos \theta_2) - c_3 \sin \theta_1 \sin \theta_2]/h^2 + c_4$  is the determinant of  $\hat{\mathbf{N}}_h(\xi, \theta)$ , where  $c_1 = 2a\sigma_{22}(\xi) + 2b\sigma_{11}(\xi)$ ,  $c_2 = 2b\sigma_{22} + 2a\sigma_{11}$ ,  $c_3 = 2c\sigma_{12}(\xi)$ , and  $c_4 = \sigma_{11}(\xi)\sigma_{22}(\xi) - \sigma_{12}^2(\xi)$ . By the second-order sufficient conditions as given in the fundamentals of optimization,  $\max_{\theta \in \Theta}\{|\det(\hat{\mathbf{N}}_h(\xi, \theta))|\}$  is obtained at  $\theta_{\max} = (\pi, \pi)$  and  $\min_{\theta \in \Theta_{\text{high}}}\{|\det(\hat{\mathbf{N}}_h(\xi, \theta))|\}$  is obtained at  $\theta_{\min} = (-\frac{1}{2}\pi, 0)$  and  $(0, -\frac{1}{2}\pi)$ . Thus  $E_h(\mathbf{N}_h(\xi))$  is bounded away from zero for all reasonable choices of  $a, \mu, \lambda$  and for all possible values of  $h > 0$ ,  $\sigma_{11}(\xi)$ ,  $\sigma_{12}(\xi)$ ,  $\sigma_{21}(\xi)$ , and  $\sigma_{22}(\xi)$  (i.e., the results do not depend on the given images and their sizes) over the whole discrete domain  $\Omega^h$ . For a very small mesh size one obtains

$$\lim_{h \rightarrow 0} E_h(\mathbf{N}_h(\xi)) = \frac{ab}{4(a+b)^2} \in (0, \frac{1}{4}).$$

As a result, it guarantees that there exists an efficient pointwise smoother within an MG method

for solving the discrete nonlinear system (5). Surprisingly, the discrete nonlinear system may contain grid anisotropies for  $\lambda \gg \mu$ . That is, the values of the coefficients referring to the different spatial directions (i.e.,  $2\bar{\mu} + \bar{\lambda}$  and  $\bar{\mu}$ ) vary considerably for such a choice of the Lamé constants. Apparently, the smoothing properties of a proper pointwise smoother for this discrete system are not affected by this scalar grid anisotropy.

### The proposed smoother and its smoothing analysis

In the previous section, the LFA shows theoretical choice of error-smoothing procedures for an MG solution of the discrete nonlinear system (5). To obtain a high-potential pointwise smoother, this work proposes an outer-inner iteration method.

Starting from an initial guess  $\mathbf{u}^{[0]}$  (typically  $\mathbf{u}^{[0]} = \mathbf{0}$ ) in the outer iteration, the proposed method computes a sequence of approximate solutions  $\mathbf{u}^{[1]}, \dots, \mathbf{u}^{[v]}, \mathbf{u}^{[v+1]}, \dots$  by solving the linearized discrete version of (5) as given by

$$\bar{\mathbf{N}}\mathbf{u}^{[v+1]} = \bar{\mathbf{G}}[\mathbf{u}^{[v]}] \quad (9)$$

where  $\bar{\mathbf{N}} = -\text{diag}(\mathcal{L}_1, \mathcal{L}_2)/h^2$  and

$$\bar{\mathbf{G}}[\mathbf{u}^{[v]}] = \begin{bmatrix} g_1 - f_1(\mathbf{u}^{[v]}) & 0 \\ 0 & g_2 - f_2(\mathbf{u}^{[v]}) \end{bmatrix} + \frac{1}{h^2} \begin{bmatrix} 0 & \mathcal{P} \\ \mathcal{P} & 0 \end{bmatrix} \begin{bmatrix} u_1^{[v]} \\ u_2^{[v]} \end{bmatrix}$$

$g_1 = g_2 = 0$  on the finest grid in the MG setting to be used shortly,  $v$  denotes the index for the outer iteration step and the  $h$  in (5) is dropped for simplicity. Next the resulting linear system is solved by the so-called block or pointwise collective Gauss-Seidel (PCGS) relaxation method for the inner iteration. Note that one may introduce a relaxation parameter  $\omega \in (0, 2)$  to obtain the  $\omega$ -PCGS (or successive over relaxation) steps. However, the LFA to be discussed shortly informs us that the smoothing properties of the PCGS relaxation method are better than those of the  $\omega$ -PCGS relaxation (SOR) method. It is also worth noting that other iterative techniques, such as the line relaxation techniques or the preconditioned conjugate gradient methods may be used as an inner solver. However, the PCGS relaxation method appears to be a cheaper option for practical applications.

To verify the smoothing efficiency of the proposed smoother, a quantitative measure is the so-called smoothing factor from an LFA which is defined as the worst asymptotic error reduction by

performing one smoothing step. To this end, let  $\tilde{\mathbf{N}}_h \mathbf{u}^h = \tilde{\mathbf{G}}(\tilde{\mathbf{u}}^h)$  be the linearized discrete system resulting from the outer iterations in (9) with  $\mathbf{u}^h$  and  $\tilde{\mathbf{u}}^h$  being the exact result and current approximation, respectively. The local inner iterations using the PCGS relaxation method leads to

$$\tilde{\mathbf{N}}_h^{[+]} \tilde{\mathbf{u}}_{\text{new}}^h + \tilde{\mathbf{N}}_h^{[0]} \tilde{\mathbf{u}}_{\text{new}}^h + \tilde{\mathbf{N}}_h^{[-]} \tilde{\mathbf{u}}_{\text{old}} = \tilde{\mathbf{G}}(\tilde{\mathbf{u}}^h), \quad (10)$$

where  $\tilde{\mathbf{u}}_{\text{new}}^h$  and  $\tilde{\mathbf{u}}_{\text{old}}^h$  stand for the approximations to  $\mathbf{u}^h$  before and after the inner smoothing step, respectively. Here

$$\begin{aligned} \tilde{\mathbf{N}}_h^{[+/-]} &= \frac{1}{h^2} \begin{bmatrix} -\mathcal{L}_1^{h[+/-]} & 0 \\ 0 & -\mathcal{L}_2^{h[+/-]} \end{bmatrix}, \\ \mathcal{L}_1^{h[+]} &\triangleq \begin{bmatrix} 0 & 0 & 0 \\ b & 0 & 0 \\ 0 & a & 0 \end{bmatrix}, \\ \mathcal{L}_1^{h[0]} &\triangleq \begin{bmatrix} 0 & 0 & 0 \\ 0 & -2(a+b) & 0 \\ 0 & 0 & 0 \end{bmatrix}, \\ \mathcal{L}_1^{h[-]} &\triangleq \begin{bmatrix} 0 & a & 0 \\ 0 & 0 & b \\ 0 & 0 & 0 \end{bmatrix}, \end{aligned}$$

$$\mathcal{L}_2^{h[+]} = (\mathcal{L}_1^{h[+]})^T, \quad \mathcal{L}_2^{h[0]} = \mathcal{L}_1^{h[0]}, \quad \text{and} \quad \mathcal{L}_2^{h[-]} = (\mathcal{L}_1^{h[-]})^T.$$

By subtracting (10) from  $\tilde{\mathbf{N}}_h \mathbf{u}^h = \tilde{\mathbf{G}}(\tilde{\mathbf{u}}^h)$ , one obtains the system of local error equations  $\tilde{\mathbf{N}}_h^{[+]} \tilde{\mathbf{e}}_{\text{new}}^h + \tilde{\mathbf{N}}_h^{[0]} \tilde{\mathbf{e}}_{\text{new}}^h + \tilde{\mathbf{N}}_h^{[-]} \tilde{\mathbf{e}}_{\text{old}} = 0$  or  $\tilde{\mathbf{e}}_{\text{new}}^h = \mathbf{S}_h \tilde{\mathbf{e}}_{\text{old}}^h$  where  $\tilde{\mathbf{e}}_{\text{new}}^h = \mathbf{u}^h - \tilde{\mathbf{u}}_{\text{new}}^h$  and  $\tilde{\mathbf{e}}_{\text{old}}^h = \mathbf{u}^h - \tilde{\mathbf{u}}_{\text{old}}^h$  are the error functions and  $\mathbf{S}_h = -[\tilde{\mathbf{N}}_h^{[0]} + \tilde{\mathbf{N}}_h^{[+]}]^{-1} [\tilde{\mathbf{N}}_h^{[-]}]$  is the amplification factor. The effect of  $\mathbf{S}_h$  on the grid functions  $\varphi_h(\theta, \mathbf{x})$  within  $\Theta_{\text{high}}$  will determine the smoothing properties of the proposed smoother. Thus the smoothing factor is defined by

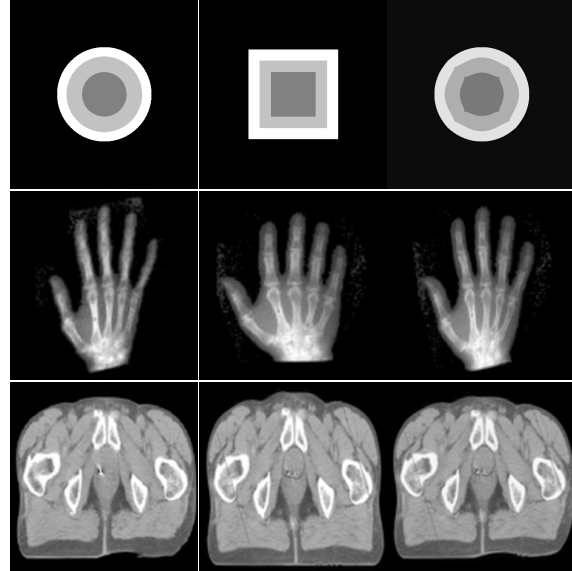
$$\tilde{\mu}_{\text{loc}} = \sup\{|\rho(\hat{\mathbf{S}}_h(\theta))| : \theta \in \Theta_{\text{high}}\}, \quad (11)$$

where  $\rho$  indicates the spectral radius of  $\hat{\mathbf{S}}_h(\theta)$ . Here  $\hat{\mathbf{S}}_h(\theta)$  is the Fourier coefficient of  $\mathbf{S}_h$  and is defined by

$$\hat{\mathbf{S}}_h(\theta) = -[\hat{\mathbf{N}}_h^{[0]}(\theta) + \hat{\mathbf{N}}_h^{[+]}(\theta)]^{-1} [\hat{\mathbf{N}}_h^{[-]}(\theta)] \in \mathbb{C}^2 \times \mathbb{C}^2, \quad (12)$$

where

$$\begin{aligned} \hat{\mathbf{N}}_h^{[+/-]}(\theta) &= \frac{1}{h^2} \begin{bmatrix} -\hat{\mathcal{L}}_1^{h[+/-]}(\theta) & 0 \\ 0 & -\hat{\mathcal{L}}_2^{h[+/-]}(\theta) \end{bmatrix}, \end{aligned}$$



**Fig. 1** Registration results for the synthetic and real medical images using the proposed MG solution. Left column: reference  $R$ ; centre column: template  $T$ ; right column: the transformed image  $T_u$  obtained from the proposed multigrid solution.

and  $\hat{\mathcal{L}}_1^{h[+]}(\theta) = a e^{-i\theta_1} + b e^{-i\theta_2}$ ,  $\hat{\mathcal{L}}_1^{h[0]}(\theta) = -2(a+b)$ ,  $\hat{\mathcal{L}}_1^{h[-]}(\theta) = b e^{i\theta_2} + a e^{i\theta_1}$ ,  $\hat{\mathcal{L}}_2^{h[+]}(\theta) = b e^{-i\theta_1} + a e^{-i\theta_2}$ ,  $\hat{\mathcal{L}}_2^{h[0]}(\theta) = -2(a+b)$ ,  $\hat{\mathcal{L}}_2^{h[-]}(\theta) = a e^{i\theta_2} + b e^{i\theta_1}$  are the Fourier coefficients of  $\mathcal{L}_1^{h[+]}$ ,  $\mathcal{L}_1^{h[0]}$ ,  $\mathcal{L}_1^{h[-]}$ ,  $\mathcal{L}_2^{h[+]}$ ,  $\mathcal{L}_2^{h[0]}$  and  $\mathcal{L}_2^{h[-]}$ , respectively. Note that the amplification factor for the local inner iterations being the  $\omega$ -PCGS or SOR relaxation method is given by

$$\begin{aligned} \hat{\mathbf{S}}_h(\theta) &= [\hat{\mathbf{N}}_h^{[0]}(\theta) + \omega \hat{\mathbf{N}}_h^{[+]}(\theta)]^{-1} \\ &\quad \times [(1-\omega) \hat{\mathbf{N}}_h^{[0]}(\theta) - \omega \hat{\mathbf{N}}_h^{[-]}(\theta)] \in \mathbb{C}^2 \times \mathbb{C}^2. \end{aligned}$$

The effectiveness of the proposed smoother is now tested by computing their smoothing factors at different values of  $\mu/\lambda$ , starting at  $10/10^{-1}$ , on a data set of two synthetic images as shown in the top row of Fig. 1 on a  $32 \times 32$  grid. Table 1 not only indicates that the proposed smoother with  $\omega = 1$  (the PCGS relaxation method) is indeed the best smoother (the smaller the smoothing factor, the better the smoothing) compared with those of the  $\omega$ -PCGS relaxation method at different values of  $\omega$  ( $\omega = 0.67, 0.75, 1.25, 1.33$ ), but also its smoothing property is not affected significantly by this ratio. Note that a small or large ratio as  $\mu/\lambda \rightarrow 0$  or  $\mu/\lambda \rightarrow \infty$  is meaningless for the purpose of image registration since it may lead to very irregular solutions.



**Table 1** Smoothing factors  $\bar{\mu}_{\text{loc}}$  by the local Fourier analysis for various  $\omega$  values after one smoothing step (one iteration) with  $\alpha = 1/10$  at different values of  $\mu/\lambda$  on a  $32 \times 32$  grid for the registration problem as shown in the top row of Fig. 1.

$\mu/\lambda$ :	10	5	1	0.5	$10^{-1}$
$\omega = 0.67$	0.7260	0.7333	0.7787	0.8179	0.9252
$\omega = 0.75$	0.6860	0.6934	0.7435	0.7876	0.9113
$\omega = 1.00$	0.5718	0.5785	0.5840	0.5901	0.6014
$\omega = 1.25$	0.5859	0.5870	0.5967	0.6168	0.7768
$\omega = 1.33$	0.6244	0.6248	0.6285	0.6344	0.7452

From a practical point of view, only moderate values of  $\mu/\lambda$ , particularly  $\mu/\lambda \in [10^{-1}, 10^1]$ , yield a reasonable deformation and lead to visually pleasing registration results. Moreover, this recommended range of  $\mu/\lambda$  provides a positive lower bound for  $E_h(\mathbf{N}_h(\xi))$  and ensures the existence of the efficient pointwise smoother. Several numerical tests given later back up the theoretical results discussed in this section; see Table 2.

### Coarsening, prolongation and restriction operators

As a starting point, the standard coarsening is used in computing the coarse-grid domain  $\Omega_H$  by doubling the grid size in each space direction, i.e.,  $h \rightarrow 2h = H$ . To fulfil the relation between the order of the differential operator and the sum of the orders of prolongation and restriction, we use the bilinear interpolation operator  $u_l^h = I_H^h u_l^H$  for prolongation and the full weighting operator  $u_l^H = I_h^H u_l^h$  for restriction.

### Coarse grid operator

Owing to the uniform and structure grids, one can use the so-called discretization coarse grid approximation (DCA) method<sup>6</sup>. The idea is to re-discretize the PDEs in (4) directly. In the case of  $f_l^H(u_1^H, u_2^H)$ , the restriction operator  $I_h^H$  is used to obtain  $u_1^H, u_2^H, R^H$  and  $T^H$ , and then compute the corresponding coarse-grid part of  $f_l^H(u_1^H, u_2^H)$ . For  $\mathcal{L}_l^H$  and  $\mathcal{P}^H$ , the corresponding coarse-grid part of  $\mathcal{L}_l^H$  and  $\mathcal{P}^H$ , are obtained using  $u_l^H$  and the DCA method. Note that it is possible to use a Galerkin operator in determining the coarse grid operators. However, a straightforward implementation requires either additional memory or additional computation for each operator application. Thus the DCA method is appropriate for numerical implementation on modest computational architecture.

**Table 2** Registration results by the proposed MG solution with  $\alpha = 1/10$ ,  $\mu = 1/2$  for Examples 1–3.

$q$	$M/D/C$		
	Example 1 $\lambda = 5/2$	Example 2 $\lambda = 1/4$	Example 3 $\lambda = 1/4$
7	8/0.0190/7.3	7/0.1168/6.2	7/0.1147/6.0
8	8/0.0146/16.9	6/0.1226/13.1	7/0.1030/14.8
9	7/0.0142/44.2	6/0.1268/37.5	6/0.1000/38.1
10	6/0.0132/129	6/0.1284/127	6/0.1002/122
11	5/0.0130/446	5/0.1289/378	6/0.1005/453
12	5/0.0128/1663	5/0.1290/1669	6/0.1006/2011

$q = -\log_2 h$ ;  $M$ : number of MG steps;  $D$ : relative reduction of dissimilarity;  $C$ : total run times (s).

### Coarsest grid solver

To overcome nonlinearity, the explicit time marching scheme associated with (5) as given by

$$\mathbf{u}_{t_{k+1}}^h = \mathbf{u}_{t_k}^h + \tau(\mathbf{G}_h[\mathbf{u}_{t_k}^h] - \mathbf{N}_h \mathbf{u}_{t_k}^h), \quad (13)$$

is used to linearize and solve the nonlinear residual equation on the coarsest grid, typically the  $4 \times 4$  grid. Here  $\tau > 0$  is the time step determined from discretization by a forward difference approximation of the time derivative  $\partial_t \mathbf{u}$ , where  $t$  is the artificial time variable.

### Multigrid cycle

The above multigrid components are linked together to form a V-, W-, or F-cycle. In this work, the W-cycle is selected for solving (5). As can be seen in the numerical experiments, the W(3,3)-cycle is very effective and can significantly decrease the relative residuals.

### AN APPLICATION FOR REGISTERING A GROUP OF MONOMODAL IMAGES

Given a finite set of monomodal images  $\mathcal{S} = \{I^1, I^2, \dots, I^N\}$  ( $N > 2$ ) where  $I^i : \Omega \subset \mathbb{R}^2 \rightarrow V \subset \mathbb{R}$  for  $i \in \{1, 2, \dots, N\}$ , the aim of the so-called implicit reference based groupwise (IRG) image registration method<sup>7</sup> is to simultaneously estimate  $N$  deformations defined by

$$\mathbf{u}^{iR} : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \mathbf{u}^{iR} : \mathbf{x} \mapsto \mathbf{u}^{iR}(\mathbf{x}) = (u_1^{iR}(\mathbf{x}), u_2^{iR}(\mathbf{x}))^T$$

from each image  $I^i \in \mathcal{S}$  to an implicit or common reference  $R$  corresponding to the group average (or the average shape). This registration problem can be stated as follows:

$$\min_{\mathbf{U}} \{ \bar{\mathcal{J}}_{\alpha}(\mathbf{U}; I^i) = \bar{\mathcal{D}}(\mathbf{U}; I^i) + \alpha \bar{\mathcal{R}}^{\text{elas}}(\mathbf{U}) \}, \quad (14)$$

where  $\mathbf{U} = (\mathbf{u}^{1R}, \mathbf{u}^{2R}, \dots, \mathbf{u}^{NR}) \in \mathcal{U}^N$  is searched over a space of admissible functions minimizing  $\bar{\mathcal{J}}_{\alpha}$ . The

first term in (14) defined by

$$\mathcal{D}(\mathbf{U}; I^i) = \frac{1}{2} \sum_{i=1}^N \sum_{i < j} \int_{\Omega} (I_{\mathbf{u}^{iR}}^i(\mathbf{x}) - I_{\mathbf{u}^{jR}}^j(\mathbf{x}))^2 d\mathbf{x}, \quad (15)$$

is used to eliminate an unavoidable bias caused by the reference selection, while the second term in (14) defined by  $\mathcal{R}^{\text{elas}}(\mathbf{U}) = \sum_{i=1}^N \mathcal{R}^{\text{elas}}(\mathbf{u}^{iR})$  is used to estimate deformations exhibiting some kind of smoothness by the elastic regularizer. Here  $I_{\mathbf{u}^{iR}}^i(\mathbf{x}) = I^i(\varphi^{iR}(\mathbf{x}))$ , where  $\varphi^{iR}(\mathbf{x}) = \mathbf{x} + \mathbf{u}^{iR}(\mathbf{x})$  denotes the transformation from image  $i$  to the implicit reference  $R$ . Note that the transformation  $\varphi^{ij}$  between every pair of images is obtained by concatenating transformations,  $\varphi^{ij}(\mathbf{x}) = \varphi^{iR}(\varphi^{jR^{-1}}(\mathbf{x}))$ . As a consequence, the minimization problem in (14) yields the system of  $N$  nonlinear PDEs. Here the  $i$ th equation is given by

$$\mathbf{f}^{iR}(\mathbf{u}^{iR}) - (\bar{\mu} \Delta \mathbf{u}^{iR} + (\bar{\lambda} + \bar{\mu}) \nabla \text{div} \mathbf{u}^{iR}) = \mathbf{0}, \quad (16)$$

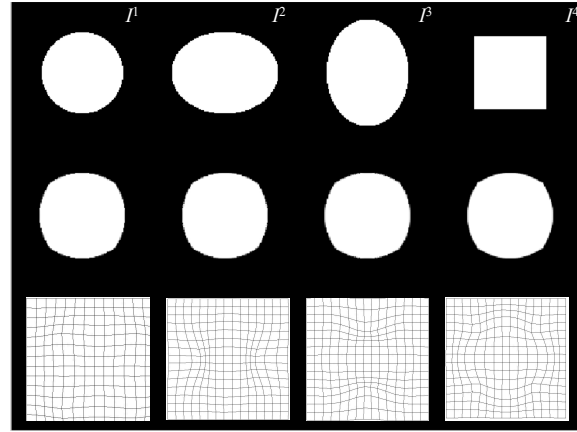
subject to the boundary conditions  $u_1^{iR} = u_2^{iR} = 0$  for all  $i \in \{1, 2, \dots, N\}$  and

$$\mathbf{f}^R(\mathbf{u}^{iR}) = \sum_{j=1}^N (I_{\mathbf{u}^{iR}}^i - I_{\mathbf{u}^{jR}}^j) \nabla_{\mathbf{u}^{iR}} I_{\mathbf{u}^{iR}}^i.$$

Apparently, the number of unknown variables in the above system can be very large and it requires an efficient solution method. To this end, the proposed MG solution can apply only a single W-cycle for each nonlinear equation. As can be seen in the next section, the proposed MG solution is very efficient at estimating the deformations and delivering the implicit reference images in a reasonable amount of time.

## NUMERICAL EXPERIMENTS

This section presents some experiments on three registration problems consisting of the synthetic and real medical images (Fig. 1) to (i) demonstrate the performance of the proposed MG solution with regard to parameter changes; (ii) show the effectiveness of the proposed MG solution in registering a group of monomodal images; (iii) compare with other MG solution methods. All numerical algorithms were implemented on a 2.5 GHz laptop with an Intel Core i5 and 8 GB of RAM. Note that in all experiments,  $\nu_1 = \nu_2 = 3$  (pre- and post-smoothing steps), the maximum number of inner iterations used by the proposed smoother is 2, the relaxation parameter  $\omega = 1$ , and cubic interpolation was used for computing the deformed template images.



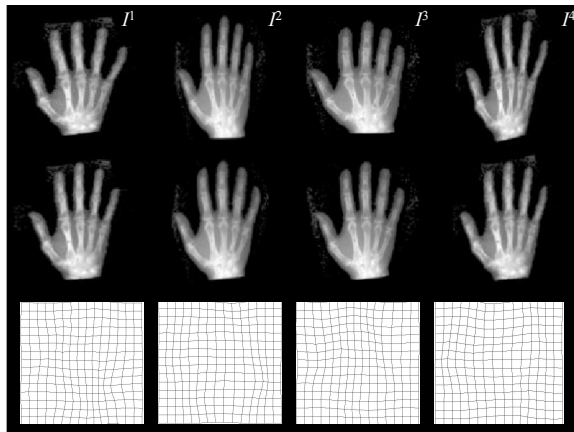
**Fig. 2** Registration results from the proposed MG solution. Top row: original four images; middle row: transformed images corresponding to the implicit reference image; bottom row: the transformation  $\varphi^{iR}$  applied to a rectangular grid.

## $h$ -independent convergence tests

One of the key properties of MG techniques is that their convergence does not depend on the number of grid points. Thus the first test was designed to investigate this property with the proposed MG solution and to support the theoretical results from LFA. The number of MG steps used to reduce the mean of the relative residual to below  $10^{-8}$ , the relative reduction of dissimilarity, and run times (in seconds) are given in Table 2 with different numbers of grid points. As can be seen in Table 2, the results show that the proposed MG solution not only converges within a few MG steps, but it is also accurate because the dissimilarities between the reference and registered images have been reduced by more than 98% for Example 1, 87% for Example 2, and 89% for Example 3.

## Tests of the proposed MG solution in registering a group of monomodal images

The aim of this test is to assess the performance of the proposed MG solution in registering a group of monomodal images. To this end, two data sets of the four synthetic and real medical images shown in the top rows of Figs. 2 and 3 were used with the parameters  $\alpha = 1/10$ ,  $\mu = 1/2$ ,  $\lambda = 1/4$ , and  $h = 1/128$ . To obtain the registration results quickly, the proposed MG solution was performed using only a single W-cycle and took about 1 s to solve each equation in (16). The results shown in the middle rows of Figs. 2 and 3 indicate that the four implicit



**Fig. 3** Registration results from the proposed MG solution. Top row: original four images; middle row: transformed images corresponding to the implicit reference image; bottom row: transformation  $\varphi^{IR}$  applied to a rectangular grid.

reference images agree with each other very well.

### Comparison with other multigrid methods

Numerical methods in Refs. 5, 8, 9 are some existing MG techniques used to solve the elastic image registration. This section uses Example 1 shown in the first two images in the top row of Fig. 1 to compare the proposed MG solution and other three MG methods by starting with the fixed parameters  $h = 1/128$ ,  $\alpha = 1/10$ ,  $\mu = 1/2$ ,  $\lambda = 5/2$  and the initial solution  $\mathbf{u}^{(0)} = \mathbf{0}$ . The linear multigrid (LMG) methods combined with the Gauss-Newton (GN) methods in Refs. 8, 9 were performed using two pre-smoothing and two post-smoothing steps until the relative residuals are below a user-supplied threshold ( $10^{-8}$ ). Table 3 reports the results for all MG methods. As expected from the experiments, all methods are very fast and accurate in registering the given images because the dissimilarities between the reference and registered images have been significantly reduced by more than 97% within the first 20 iterations. The proposed method is not only the fastest way to solve this registration problem, but also reduces the relative residuals to below  $10^{-8}$ .

### CONCLUSIONS

This paper has presented an efficient multigrid solution of the nonlinear PDEs arising in elastic image registration. The  $h$ -ellipticity shows that the resulting nonlinear discrete operator is amenable to multigrid methods. We propose a high potential smoother based on an outer-inner iteration scheme

**Table 3** A comparison among different multigrid methods in Refs. 5, 8, 9 to solve the elastic image registration in the first 20 iterations.

Methods	$M/R/D$
FAS-NMG V(1,1) <sup>5</sup>	$20/2.7 \times 10^{-6}/0.0196$
GN+LMG V(2,2) <sup>8</sup>	$20/4.1 \times 10^{-7}/0.0214$
GN+LMG F(2,2) <sup>9</sup>	$20/7.5 \times 10^{-8}/0.0211$
Proposed method F(2,2)	$14/8.9 \times 10^{-9}/0.0217$
Proposed method W(3,3)	$8/1.7 \times 10^{-9}/0.0189$

$M$ : number of iterations or MG steps to reduce the relative residuals to below  $10^{-8}$ ;  $R$ : relative residuals;  $D$ : relative reduction of dissimilarity.

and confirm its effectiveness by the LFA and several numerical tests conducted to assess the accuracy and efficiency of the proposed multigrid solution. As expected from a multigrid method, numerical results show that the proposed multigrid solution is reliable and can be recommended for a wide range of real applications.

**Acknowledgements:** The authors express their thanks to the referees. The first author's work was partially supported by the Thailand Research Fund Grant TRG5680037 and the Centre of Excellence in Mathematics.

### REFERENCES

1. Fischer B, Modersitzki J (2002) Fast diffusion registration. *Contemp Math* **313**, 117–29.
2. Fischer B, Modersitzki J (2003) Curvature based image registration. *J Math Imag Vis* **18**, 81–5.
3. Chumchob N, Chen K, Brito C (2011) A fourth order variational image registration model and its fast multigrid algorithm. *Multiscale Model Simulat* **9**, 89–128.
4. F-Schauf C, Henn S, Witsch K (2008) Multigrid based total variation image registration. *Comput Visual Sci* **11**, 101–13.
5. Henn S, Witsch K (2001) Iterative multigrid regularization techniques for image matching. *SIAM J Sci Comput* **23**, 1077–93.
6. Trottenberg U, Oosterlee C, Schuller A (2001) *Multigrid*, Academic Press.
7. Geng X, Christensen G, Gu H, Ross T, Yang Y (2009) Implicit reference-based group-wise image registration and its application to structural and functional MRI. *NeuroImage* **47**, 1341–51.
8. Haber E, Modersitzki J (2006) A multilevel method for image registration. *SIAM J Sci Comput* **27**, 1594–607.
9. Hömke L (2006) A multigrid method for anisotropic PDEs in elastic image registration. *Numer Lin Algebra Appl* **13**, 215–29.