Improved Variational Image Registration Model and a Fast Algorithm for Its Numerical Approximation

Noppadol Chumchob,[1,2*] **Ke Chen**[3*]
[1] *Department of Mathematics, Faculty of Science, Silpakorn University, Nakorn Pathom 73000, Thailand*

[2] *Centre of Excellence in Mathematics, CHE, Bangkok 10400, Thailand*

[3] *Department of Mathematical Sciences, The University of Liverpool, Liverpool L69 7ZL, United Kingdom*

Received 4 October 2010; accepted 10 August 2011
Published online 2 November 2011 in Wiley Online Library (wileyonlinelibrary.com).
DOI 10.1002/num.20710

In a multimodal image registration scenario, where two given images have similar features, but noncomparable intensity variations, the sum of squared differences is not suitable for inferring image similarities. In this article, we first propose a new variational model based on combining intensity and geometric transformations, as an alternative to use mutual information and an improvement to the work by Modersitzki and Wirtz (Modersitzki and Wirtz, Lect Notes Comput Sci 4057 (2006), 257–263), and then develop a fast multigrid (MG) algorithm for solving the underlying system of fourth-order and nonlinear partial differential equations. We can demonstrate the effective smoothing property of the adopted primal-dual smoother by a local Fourier analysis. Numerical tests will be presented to show both the improvements achieved in image registration quality and MG efficiency. © 2011 Wiley Periodicals, Inc. Numer Methods Partial Differential Eq 28: 1966–1995, 2012

Keywords: curvature image registration; deformable image registration; multimodal image registration; nonlinear multigrid; nonrigid image registration; variational image registration

I. INTRODUCTION

Under many real-world conditions, even intensity variations of two given images taken from the same object on the same scanner within the same protocol can be locally or globally different, for example, clinical magnetic resonance (MR) images affected by the signal intensity inhomogeneity (bias field) due to imperfections in the radiofrequency coils and object-dependent interactions [1–5]. In general, registration of different image modalities is a common, but demanding task. The

Correspondence to: Ke Chen, Department of Mathematical Sciences, The University of Liverpool, Peach Street, Liverpool L69 7ZL, United Kingdom (e-mail: k.chen@liv.ac.uk)
*Both authors are members of the Centre for Mathematical Imaging Techniques.

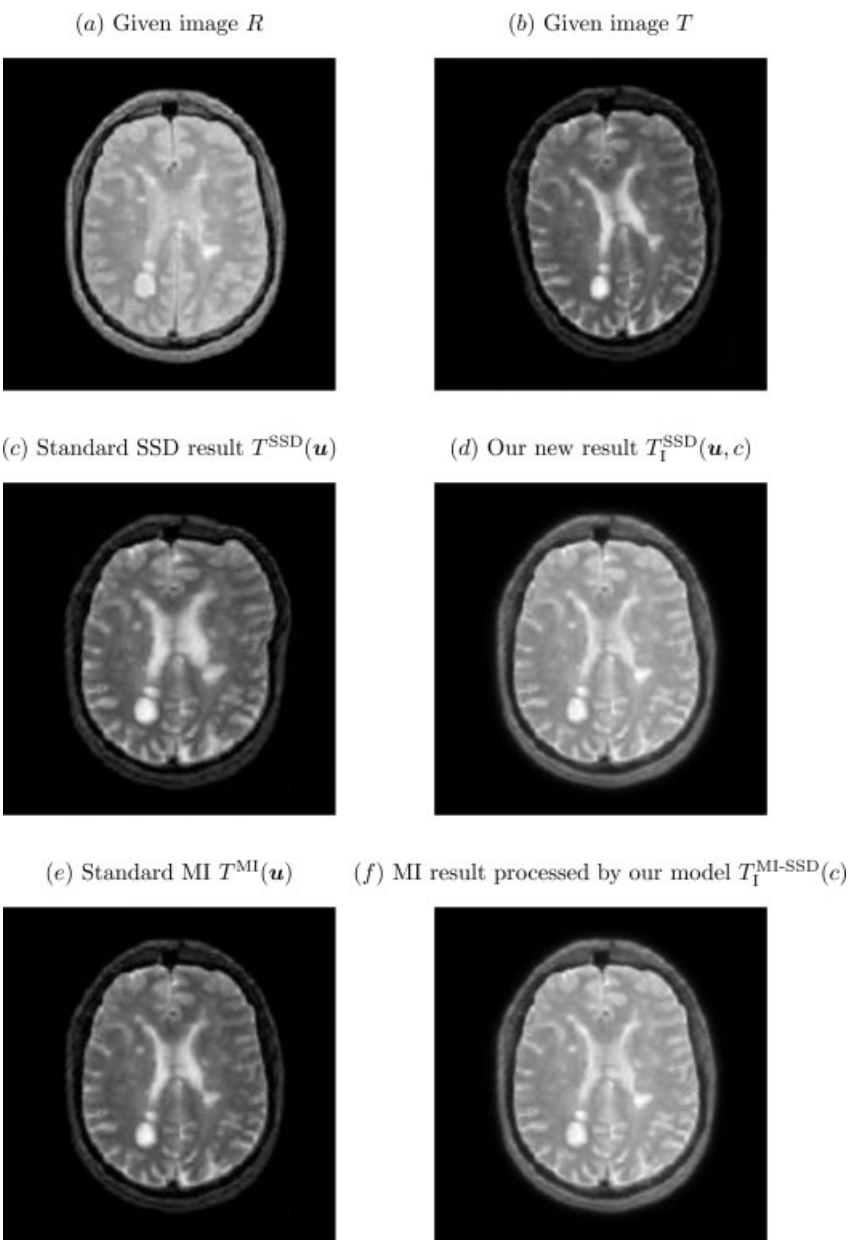

FIG. 1. Comparison of numerical results by SSD and MI with our new model. Top row: a registration problem consisting a pair of MR images of a human head shown in (a) reference R and (b) template T. Middle row: two registered images (c) $T^{\mathrm{SSD}}(\boldsymbol{u})$ by SSD (poor) and (d) the improved result $T_{\mathrm{I}}^{\mathrm{SSD}}(\boldsymbol{u}, c)$ by our proposed variational model (14). Bottom row: two registered images (e) $T^{\mathrm{MI}}(\boldsymbol{u})$ by MI (poor) and (f) the improved result $T_{\mathrm{I}}^{\mathrm{MI-SSD}}(c)$ by our proposed variational model (17) as a standardization between R and $T^{\mathrm{MI}}(\boldsymbol{u})$. Notice that first our new model (14) accurately registers the images without any additional preprocessing steps, and secondly the model (17) is effective in normalizing (postprocessing) the intensity variations between the images.

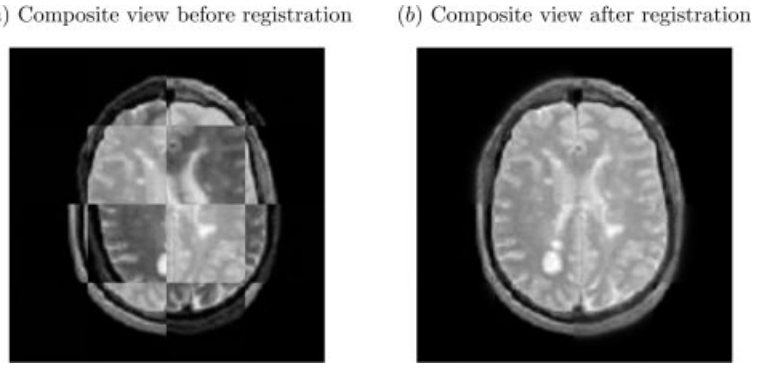

FIG. 2. Composite views between the images before and after registration for the problem shown in Fig. 1(a, b). (a) Drastically different composite view between R and T before registration; (b) indistinguishable composite view between R and $T_{\mathrm{I}}^{\mathrm{SSD}}(\boldsymbol{u})$ after registration based on our new variational model (14). The intensity variations in (b) between the images are well-matched.

sum of squared differences (SSDs) without any preprocessing steps (e.g., the intensity normalization or standardization methods) is not suitable for determining image similarities of multimodality as it reduces accuracy and efficiency of an expected registration, as illustrated in Figs. 1 and 2. Conversely, mutual information (MI) as a similarity measure is more appropriate and invariant to overall intensity scale differences. MI is often adopted to deal with the lack of a model of intensity transformations. However, it has a number of well-known drawbacks. First, MI is known to be highly nonconvex as well as nonlinear and has typically many local minima. Therefore, nonlinearity of the registration problem is enhanced by the use of MI. Second, the computation of MI and its first variation require approximations of the joint density, which summarizes the co-occurrence of events from the image intensities obtained from the given images. Such approximations are usually expensive and sensitive to some parameters, such as the width of the Parzen-window kernel and the set of local intensity samples. Finally, because of the above difficulties, there is not a unique or even common implementation for estimating MI and its first variation; see more discussions in Refs. [6–12] and references therein.

Several related methods are proposed in the literature to improve the SSD model. This article follows closely and improves on the particular work by Ref. [13], which uses a nonparametric intensity transformation for the elastic registration of monomodal images and applies the total variation (TV) energy to constraint the intensity transformations. Our improvements are three-fold. First, instead of an multiplicative correction framework, we consider a new additive correction framework. Second, for both frameworks, we propose a new curvature regularizer for the displacement field as well as the correction field. Finally, we develop a fast nonlinear multigrid (NMG) algorithm using the primal-dual fixed-point (PDFP) smoother.

Some registration models related to Ref. [13] and the current work can be found in Refs. [14–17]. In Ref. [14], the polynomial-based intensity transformation is used for elastic registration, with an iterative scheme that alternates between estimating the coefficients of the polynomial and searching the nonparametric transformation minimizing the energy functional using the demons method [18]. These coefficients have the purpose of estimating the intensity changes that match the intensity values between the images. In Ref. [15], a locally linear intensity transformation with a smoothness constraint on the contrast/brightness parameters is used to model the elastic registration of multimodal images based on a locally affine transformation with a global smoothness

constraint in a differential multiscale framework. In Ref. [16], the registration of multimodal images is modeled by a low-order polynomial intensity transformation and a global affine transformation. In Ref. [17], the registration of multimodal images is modeled using a probabilistic formulation in a multiscale framework. The main aim is to simultaneously determine the local parameters of the geometric transformation using the B-spline models by Ref. [19] and the local coefficients of the polynomial intensity transformation that lead to successful registration. These parameters and coefficients are represented as Markov random fields giving a priori information about the homogeneity of the intensity and geometric changes.

For variational models, fast methods are necessary for high-resolution digital images. For a nonlinear system like (19; see Section III), the use of a NMG method is a natural choice but its convergence is not automatic. For examples of successful developments in various image processing applications, refer to Refs. [20–33] where the Euler–Lagrange systems of partial differential equations (PDEs) can be of second- or fourth-order. Previous work on NMG techniques for deformable image registration using free-form or nonrigid deformations in Refs. [25, 27, 28, 31, 33] considers different deformation models or different multigrid (MG) components. In Ref. [25], the efficient NMG methods based on the typical FP iteration method for overcoming the singular Neumann boundary problems of discrete systems are presented for the diffusion- and modified TV-based image registration, respectively. In Ref. [27], a special treatment for singular systems due to the Neumann boundary conditions before and during the NMG method for TV-based image registration is introduced in solving the minimization problem of the SSD functional. In Ref. [28], a full-multigrid (FMG) method based on the Newton–Gauss–Seidel smoother and an adaptive smoothing approach for the deformation field are developed in the context of diffusion image registration. In Ref. [31], a NMG method based on the discretized optimality conditions for elastic image registration is presented. In Ref. [33], a FMG method based on the FP type of smoothers is developed for diffusion image registration subject to Dirichlet boundary conditions. Although the above reviewed NMG techniques have been used for other models of nonrigid image registration, to the best of our knowledge, these past methods do not work for the more challenging system of nonlinear and high-order PDEs like (19) for simultaneously determining the intensity and geometric transformations, and our proposed method is new.

The rest of this article is organized as follows. The variational image registration model of Ref. [13] and our new model combining the intensity and geometric transformations are introduced in Section II, followed by discussion of solving the resulting fourth-order PDEs by a primal-dual (PD) formulation in Section III. Section IV discusses the numerical implementation and the numerical solution for the PD formulation, in particular a multilevel approach based on an efficient NMG algorithm. The robustness of the proposed registration model and its numerical approach are illustrated using examples including clinical data in Section V. The last section is devoted to conclusions.

II. IMPROVED VARIATIONAL IMAGE REGISTRATION MODEL

Below we shall first introduce the image registration problem and the standard variational model, and then discuss the method of Ref. [13] and our proposed new formulation in the same spirit.

Suppose that two images R (the reference) and T (the template), intended for registration, are given as the continuous functions mapping from an image domain $\Omega \subset \mathbb{R}^d$ into $V \subset \mathbb{R}_0^+$ and each component u_d of $\boldsymbol{u}$ is the function of the spatial position $\mathbf{x} = (x_1, x_2, \ldots, x_d)^\top \in \Omega$. Without loss of generality, we assume that the registration problem is described in the two-dimensional (2D) case ($d = 2$) throughout this article, but it is readily extendable to the three-dimensional

case ($d = 3$). We also assume further that $\Omega = [0, 1]^2 \subset \mathbb{R}^2$ and $V = [0, 1]$ for 2D gray intensity images.

A general framework of the registration problem of monomodal images can be formulated as follows: For the given R and T, we search simultaneously for a vector-valued nonparametric transformation φ defined by

$$\varphi(\boldsymbol{u})(\cdot) : \mathbb{R}^d \to \mathbb{R}^d, \qquad \varphi(\boldsymbol{u})(\mathbf{x}) : \mathbf{x} \mapsto \mathbf{x} + \boldsymbol{u}(\mathbf{x}) \tag{1}$$

that depends on an unknown deformation or displacement field

$$\boldsymbol{u} : \mathbb{R}^d \to \mathbb{R}^d, \qquad \boldsymbol{u} : \mathbf{x} \mapsto \boldsymbol{u}(\mathbf{x}) = (u_1(\mathbf{x}), u_2(\mathbf{x}), \ldots, u_d(\mathbf{x}))^\top \tag{2}$$

and an intensity transformation f such that the transformed template

$$f(T \circ \varphi(\boldsymbol{u}(\mathbf{x}))) = f(T(\mathbf{x} + \boldsymbol{u}(\mathbf{x}))) = f(T(\boldsymbol{u}))$$

becomes similar to the reference R in a geometric sense, that is

$$R(\mathbf{x}) = f(T(\boldsymbol{u})) + \eta(\mathbf{x}). \tag{3}$$

Here, $\eta(\mathbf{x})$ is an unknown random function of uncorrelated noise. If the intensity variations of R and T are comparable, the intensity transformation f can be represented by the identity function and we search only the deformation field $\boldsymbol{u}$. Then, the common choice of the SSD functional $\mathcal{D}$ defined by

$$\mathcal{D}(\boldsymbol{u}) = \mathcal{D}^{\text{SSD}}(\boldsymbol{u}, c) = \frac{1}{2} \int_\Omega (T(\mathbf{x} + \boldsymbol{u}(\mathbf{x})) - R(\mathbf{x}))^2 d\mathbf{x}$$

is enough. Finally, the inverse problem of finding $\boldsymbol{u}$ is solved by the Tikhonov variational formulation

$$\min_{\boldsymbol{u}} \{\mathcal{J}_\alpha(\boldsymbol{u}, c) = \mathcal{D}(\boldsymbol{u}) + \alpha \mathcal{R}(\boldsymbol{u})\} \tag{4}$$

where $\mathcal{R}$ is a suitable regularizer such as one of (10)–(13). Refer to Ref. [34].

However, in real-life applications (such as MR and CT medical imaging), inhomogeneity of image intensities and noise are represented in both R, T images or one of them. If so, the above framework (4) becomes ineffective and we require a more complex intensity function f than those of parametric intensity transformations (e.g, polynomials). To design a general-purpose registration model for these cases, we assume that model f is nonparametric and obeys with the following intensity relationships

$$\text{Intensity model I: } f(T(\boldsymbol{u})) = T(\boldsymbol{u}) + c(\mathbf{x}) \quad \text{(additive intensity correction model)} \tag{5}$$

$$\text{Intensity model II: } f(T(\boldsymbol{u})) = c(\mathbf{x})T(\boldsymbol{u}) \quad \text{(multiplicative intensity correction model)} \tag{6}$$

where $c : \Omega \to \mathbb{R}$ is an unknown nonparametric intensity correction. In a previous work [13], coinciding with intensity model II, the following variational model is proposed

$$\min_{\boldsymbol{u}, c} \{\mathcal{J}_{\alpha_1, \alpha_2}(\boldsymbol{u}, c) = \mathcal{D}(\boldsymbol{u}, c) + \alpha_1 \mathcal{R}_1(\boldsymbol{u}) + \alpha_2 \mathcal{R}_2(c)\} \tag{7}$$

with

$$\mathcal{D}(\boldsymbol{u},c)=\mathcal{D}_{\mathrm{II}}^{\mathrm{SSD}}(\boldsymbol{u},c)=\frac{1}{2}\int_{\Omega}(c(\mathbf{x})T(\mathbf{x}+\boldsymbol{u}(\mathbf{x}))-R(\mathbf{x}))^2d\mathbf{x},$$

$$\mathcal{R}_1=\mathcal{R}^{\mathrm{elas}}(\boldsymbol{u})=\int_{\Omega}\left((\mu/4)\sum_{l,m=1}^{2}(\partial_{x_l}u_m+\partial_{x_m}u_l)^2+(\lambda/2)(\nabla\cdot\boldsymbol{u})^2\right)d\mathbf{x},$$

$$\mathcal{R}_2=\mathcal{R}^{L_2}(c)=\int_{\Omega}|\nabla c|^2d\mathbf{x}=\int_{\Omega}\left(c_{x_1}^2+c_{x_2}^2\right)d\mathbf{x}.$$

Further, the resulting Euler–Lagrange equations take the form

$$\begin{cases}f_1(\boldsymbol{u},c)-\alpha_1((\lambda+2\mu)\partial_{x_1x_1}u_1+\mu\partial_{x_2x_2}u_1+(\lambda+\mu)\partial_{x_1x_2}u_2)=0\\ f_2(\boldsymbol{u},c)-\alpha_1((\lambda+\mu)\partial_{x_1x_2}u_1+\mu\partial_{x_1x_1}u_2+(\lambda+2\mu)\partial_{x_2x_2}u_2)=0\\ f_3(\boldsymbol{u},c)-\alpha_2\nabla\cdot\dfrac{\nabla c}{|\nabla c|_{\beta_2}}=0\end{cases}\tag{8}$$

subject to the boundary conditions

$$\langle\mu(\nabla\boldsymbol{u}+(\nabla\boldsymbol{u})^{\top})+\lambda\mathrm{diag}(\nabla\cdot\boldsymbol{u}),\boldsymbol{n}\rangle_{\mathbb{R}^2}=0\quad\text{and}\quad\langle\nabla c,\boldsymbol{n}\rangle_{\mathbb{R}^2}=0\text{ on }\partial\Omega.\tag{9}$$

Although the above choice of regularizers in (7) that is, Ref. [13] for combining homogenization and registration is useful to some problems, the regularizer $\mathcal{R}^{\mathrm{elas}}(\boldsymbol{u})$ restricts $\boldsymbol{u}$ to be extremely smooth and to have small displacements, and the regularizer $\mathcal{R}^{L_2}(c)$ also admits smooth correction c only so the model (7) does not have generality. There are alternative regularizers one could consider such as

$$\mathcal{R}^{\mathrm{diff}}(\boldsymbol{u})=\frac{1}{2}\sum_{l=1}^{2}\int_{\Omega}|\nabla u_l|^2d\mathbf{x},\tag{10}$$

$$\mathcal{R}^{\mathrm{FMcurv}}(\boldsymbol{u})=\frac{1}{2}\sum_{l=1}^{2}(\Delta u_l)^2d\mathbf{x},\tag{11}$$

$$\mathcal{R}^{\mathrm{HWcurv}}(\boldsymbol{u})=\frac{1}{2}\sum_{l=1}^{2}\int_{\Omega}(\Delta u_l)^2-2\left(u_{l_{x_1x_1}}u_{l_{x_2x_2}}-u_{l_{x_1x_2}}^2\right)d\mathbf{x},\tag{12}$$

$$\mathcal{R}^{\beta\mathrm{TV}}(\boldsymbol{u})=\sum_{l=1}^{2}\int_{\Omega}|\nabla u_l|_{\beta_1}d\mathbf{x}=\sum_{l=1}^{2}\int_{\Omega}\sqrt{u_{l_{x_1}}^2+u_{l_{x_2}}^2+\beta_1}d\mathbf{x},\tag{13}$$

for both $\boldsymbol{u}$ and c. These regularizers (10)–(13), similarly having no generality, cannot work well for both smooth and nonsmooth registration problems.

To fundamentally improve on (7), motivated by Refs. [35–37], we propose to regularize curvatures of both $\boldsymbol{u}$ and c as they are found to be able to cope well with both smooth and nonsmooth registration problems. Thus, our new registration model is

$$\min_{\boldsymbol{u},c}\left\{\mathcal{J}_{\alpha_1,\alpha_2}(\boldsymbol{u},c)=\mathcal{D}(\boldsymbol{u},c)+\alpha_1\sum_{l=1}^{2}\int_{\Omega}\Phi(\kappa(u_l))d\mathbf{x}+\alpha_2\int_{\Omega}\Phi(\kappa(c))d\mathbf{x}\right\}\tag{14}$$

where $\alpha_1, \alpha_2 > 0$ are the regularization parameters. Here, we have chosen

$$\mathcal{R}_1(\boldsymbol{u}) = \mathcal{R}^{\text{NewCv}}(\boldsymbol{u}) = \sum_{l=1}^{2} \int_\Omega \Phi(\kappa(u_l)) d\mathbf{x}, \qquad \mathcal{R}_2(c) = \mathcal{K}(c) = \int_\Omega \Phi(\kappa(c)) d\mathbf{x},$$

the first SSD term including both intensity models I and II in (5) and (6) that is

$$\mathcal{D}(\boldsymbol{u}, c) = \begin{cases} \mathcal{D}_{\text{I}}^{\text{SSD}}(\boldsymbol{u}, c) = \dfrac{1}{2} \displaystyle\int_\Omega (T(\mathbf{x} + \boldsymbol{u}(\mathbf{x})) + c(\mathbf{x}) - R(\mathbf{x}))^2 d\mathbf{x}, & \text{Model (5)} \\ \mathcal{D}_{\text{II}}^{\text{SSD}}(\boldsymbol{u}, c) = \dfrac{1}{2} \displaystyle\int_\Omega (c(\mathbf{x}) T(\mathbf{x} + \boldsymbol{u}(\mathbf{x})) - R(\mathbf{x}))^2 d\mathbf{x}, & \text{Model (6)} \end{cases} \tag{15}$$

and $\Phi(s) = \frac{1}{2} s^2$ using the mean curvature

$$\kappa(u_l) = \nabla \cdot \frac{\nabla u_l}{|\nabla u_l|_{\beta_1}} = \frac{\left(\beta_1 + u_{l_{x_1}}^2\right) u_{l_{x_1 x_1}} - 2 u_{l_{x_1}} u_{l_{x_2}} u_{l_{x_1 x_2}} + \left(\beta_1 + u_{l_{x_2}}^2\right) u_{l_{x_2 x_2}}}{\left(\beta_1 + u_{l_{x_1}}^2 + u_{l_{x_2}}^2\right)^{3/2}}, \quad \beta_1 > 0. \tag{16}$$

Our new variational framework aims to achieve two aims: (i) it does not require an affine pre-registration step and (ii) it is more flexible for both smooth and nonsmooth registration problems than previous work [13].

Regarding the choice of a regularizer for c in (14), we note the following:

1. For a linear intensity correction $\hat{c}(\mathbf{x}) = a_1 x_1 + a_2 x_2 + a_3$, $\mathcal{K}(\hat{c}(\mathbf{x})) = 0$, that is, the nontrivial kernel of $\mathcal{K}$ consists only of the linear transformations, and consequently this energy is invariant under globally and locally linear intensity corrections. Now comparing with the previously tested choices for c in Ref. [13] namely

$$\mathcal{TV}(c) = \int_\Omega |\nabla c|_{\beta_2} d\mathbf{x} = \int_\Omega \sqrt{c_{x_1}^2 + c_{x_2}^2 + \beta_2} d\mathbf{x}, \quad \beta_2 > 0,$$
$$\mathcal{R}^{L_2}(c) = \int_\Omega |\nabla c|^2 d\mathbf{x} = \int_\Omega \left(c_{x_1}^2 + c_{x_2}^2\right) d\mathbf{x}$$

   we see that both $\mathcal{TV}(\hat{c}(\mathbf{x})) = 0$ and $\mathcal{R}^{L_2}(\hat{c}(\mathbf{x}) = 0$ hold if and only if $a_1 = a_2 = a_3 = 0$. This means that both $\mathcal{TV}$ and $\mathcal{R}^{L_2}$ do not allow nontrivial linear intensity corrections; see Fig. 7.
2. $\mathcal{K}$ preserves discontinuities of c because the diffusion coefficients of the Euler–Lagrange equations resulting from (14) are zero in regions representing large gradient of c, that is, $1/|\nabla c|_{\beta_2} \to 0$ and $\nabla c \cdot \nabla \Psi'(\kappa(c))/|\nabla c|_{\beta_2}^3 \to 0$ when $|\nabla c|_{\beta_2} \to \infty$; see (19) and (22) in Section III later. As a result, the corrected images by the regularizer $\mathcal{K}$ are not blurred and are different from those by $\mathcal{R}^{L_2}$; see Figs. 3 and 5.

The above theoretical remarks can be tested through two registration problems with their numerical results shown in Figs. 3–7. Clearly, the combination of $\mathcal{D}_{\text{I}}^{\text{SSD}}$ and $\mathcal{K}(c)$ is more suitable; as mentioned, in a previous work using $\mathcal{D}_{\text{II}}^{\text{SSD}}$ [13], only $\mathcal{R}_1(\boldsymbol{u}) = \mathcal{R}^{\text{elas}}(\boldsymbol{u})$ and $\mathcal{R}_2(c) = \mathcal{TV}(c)$ were tested.

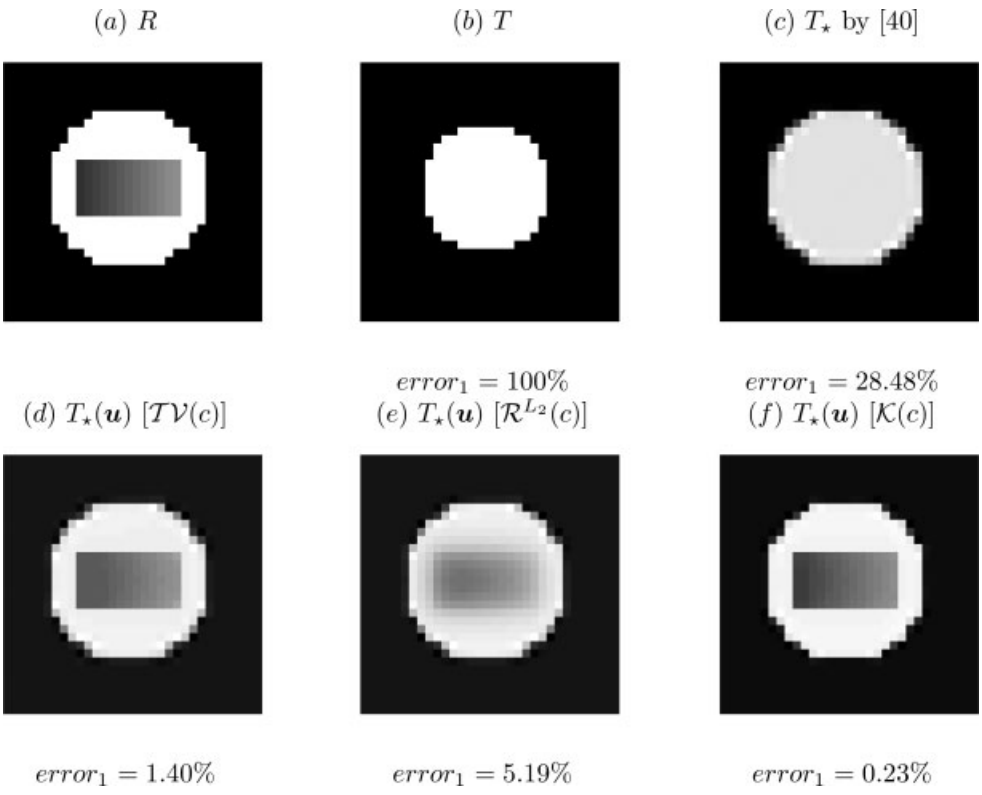

FIG. 3. A numerical test using $\mathcal{D}_1^{\mathrm{SSD}}$ (14) with three regularization techniques for c to show that our technique $\mathcal{K}(c)$ is an improvement over $\mathcal{TV}(c)$ and $\mathcal{R}^{L_2}(c)$. Top row from left to right: (a) reference R, (b) template T, and (c) registered image by Ref. [16]. Bottom row from left to right: three registered images $T_\star(\boldsymbol{u})$ via $\mathcal{TV}(c)$, $\mathcal{R}^{L_2}(c)$, and $\mathcal{K}(c)$, respectively. Clearly, the registration technique based on the polynomial intensity functions by Ref. [16] is less efficient. Here, $error_1$ denotes the percentage error.

Finally, we have two additional remarks:

Remark 1.

1. If we set $\alpha_1 = 0$, $\boldsymbol{u} = 0$ and $\alpha_2 > 0$, the following minimization problem

$$\min_c \{\mathcal{J}_{\alpha_2}(0, c) = \mathcal{D}(0, c) + \alpha_2 \mathcal{R}_2(c)\} \tag{17}$$

   gives only the nonparametric intensity correction c for the normalization or standardization between the images, that is, $T(\mathbf{x}) + c(\mathbf{x})$, $c(\mathbf{x})T(\mathbf{x}) \approx R(\mathbf{x})$; see, for example, the standardization between R and $T^{\mathrm{MI}}(\boldsymbol{u})$ using (17) in Fig. 1(e). That is to say, our reduced model (17) can be used for standardization purposes, for example, for postprocessing MI results.
2. The new variational model (14) can be adapted to solve problems related to optical flow computation or stereo disparity estimation, for example, by introducing the energy functional $\mathcal{R}_2$ in the variational formulation of optical flow computation.

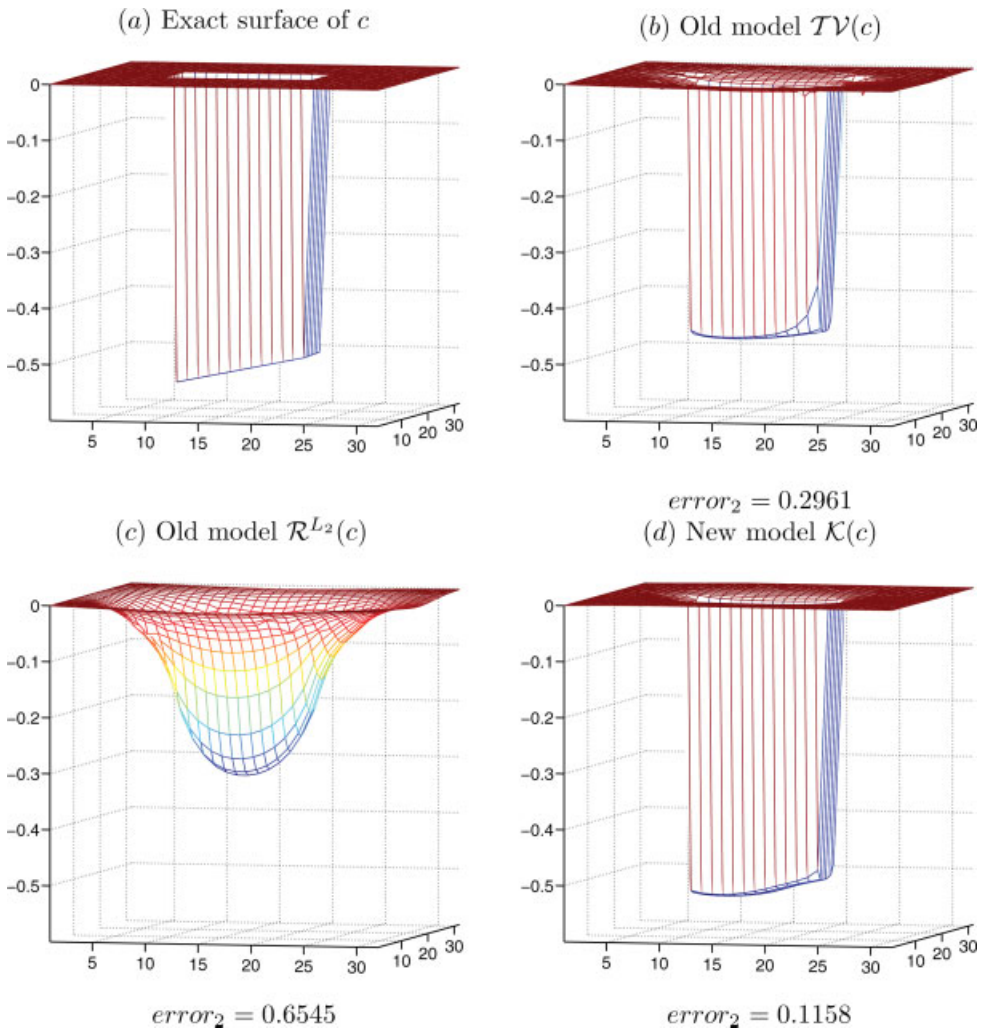

FIG. 4. Surface plots of c for the registration problem in Fig. 3(a, b). (a) the exact surface of c; (b, d) the results via $\mathcal{TV}(c)$, $\mathcal{R}^{L_2}(c)$, and $\mathcal{K}(c)$, respectively. Here, the error_2 denotes the two-norm of the differences between the exact and approximate solutions. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

III. EULER–LAGRANGE EQUATIONS AND THEIR PD FORMULATIONS

Consider our first and preferred case in (14), that is, $\mathcal{D} = \mathcal{D}_1^{\text{SSD}}$, $\mathcal{R}_1(\boldsymbol{u}) = \mathcal{R}^{\text{NewCv}}(\boldsymbol{u})$ and $\mathcal{R}_2(c) = \mathcal{K}(c)$ or

$$\min_{\boldsymbol{u},c} \left\{ \mathcal{J}_{\alpha_1,\alpha_2}(\boldsymbol{u},c) = \frac{1}{2}\int_\Omega (T(\mathbf{x}+\boldsymbol{u}(\mathbf{x})) + c(\mathbf{x}) - R(\mathbf{x}))^2 d\mathbf{x} + \alpha_1 \sum_{l=1}^{2} \int_\Omega \Phi(\kappa(u_l)) d\mathbf{x} + \alpha_2 \int_\Omega \Phi(\kappa(c)) d\mathbf{x} \right\}. \quad (18)$$

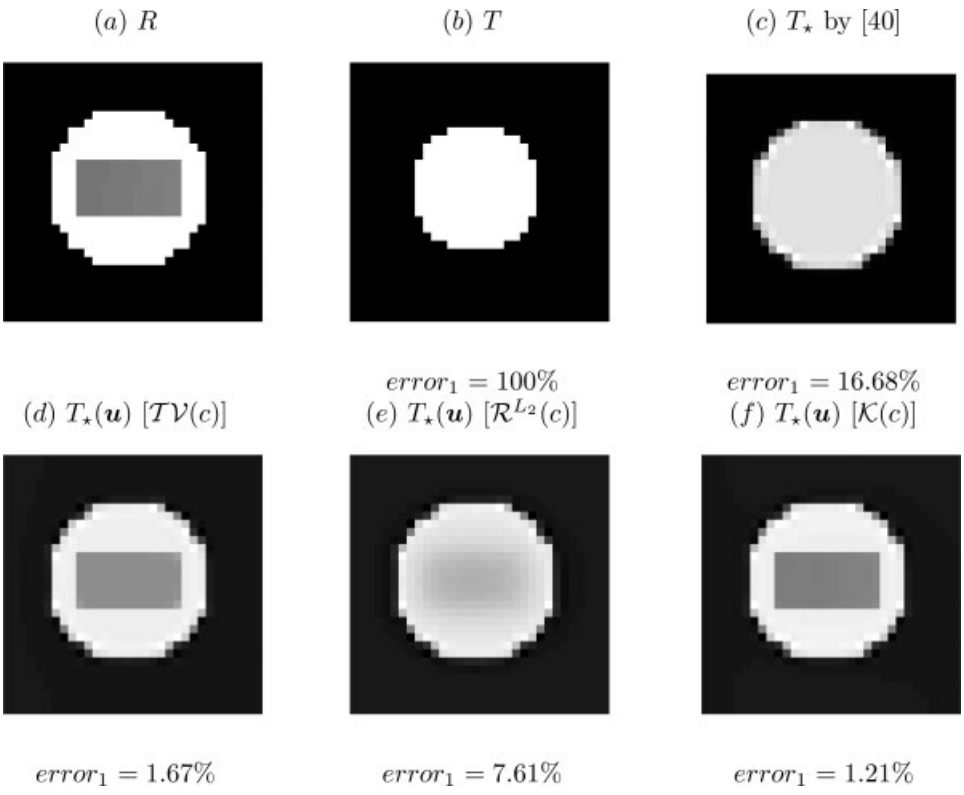

FIG. 5. A numerical test using $\mathcal{D}_{\mathrm{II}}^{\mathrm{SSD}}$ (15) with three regularization techniques for c to show that our technique $\mathcal{K}(c)$ is an improvement over $\mathcal{TV}(c)$ and $\mathcal{R}^{L_2}(c)$. Top row from left to right: (a) reference R, (b) template T, and (c) registered image via Ref. [16]. Bottom row from left to right: three registered images $T_\star(\boldsymbol{u})$ via $\mathcal{TV}(c)$, $\mathcal{R}^{L_2}(c)$, and $\mathcal{K}(c)$, respectively. Clearly, the registration technique based on the polynomial intensity functions by Ref. [16] is less efficient. Here, error_1 denotes the percentage error.

The resulting Euler–Lagrange equations of this proposed variational image registration model is given by

$$\begin{cases} \underbrace{(T(\boldsymbol{u})+c-R)\partial_{u_1}T(\boldsymbol{u})}_{f_1(\boldsymbol{u},c)} + \alpha_1\nabla\cdot\left(\dfrac{1}{|\nabla u_1|_{\beta_1}}\nabla\Phi'(\kappa(u_1)) - \dfrac{\nabla u_1\cdot\nabla\Phi'(\kappa(u_1))}{(|\nabla u_1|_{\beta_1})^3}\nabla u_1\right) = 0 \\ \underbrace{(T(\boldsymbol{u})+c-R)\partial_{u_2}T(\boldsymbol{u})}_{f_2(\boldsymbol{u},c)} + \alpha_1\nabla\cdot\left(\dfrac{1}{|\nabla u_2|_{\beta_1}}\nabla\Phi'(\kappa(u_2)) - \dfrac{\nabla u_2\cdot\nabla\Phi'(\kappa(u_2))}{(|\nabla u_2|_{\beta_1})^3}\nabla u_2\right) = 0 \\ \underbrace{(T(\boldsymbol{u})+c-R)}_{f_3(\boldsymbol{u},c)} + \alpha_2\nabla\cdot\left(\dfrac{1}{|\nabla c|_{\beta_2}}\nabla\Psi'(\kappa(c)) - \dfrac{\nabla c\cdot\nabla\Psi'(\kappa(c))}{(|\nabla c|_{\beta_2})^3}\nabla c\right) = 0 \end{cases} \tag{19}$$

subject to the boundary conditions

$$\langle\nabla u_l,\boldsymbol{n}\rangle_{\mathbb{R}^2} = 0, \qquad \langle\nabla\Phi'(\kappa(u_l)),\boldsymbol{n}\rangle_{\mathbb{R}^2} = 0 \quad \text{for } l=1,2 \text{ on } \partial\Omega \tag{20}$$

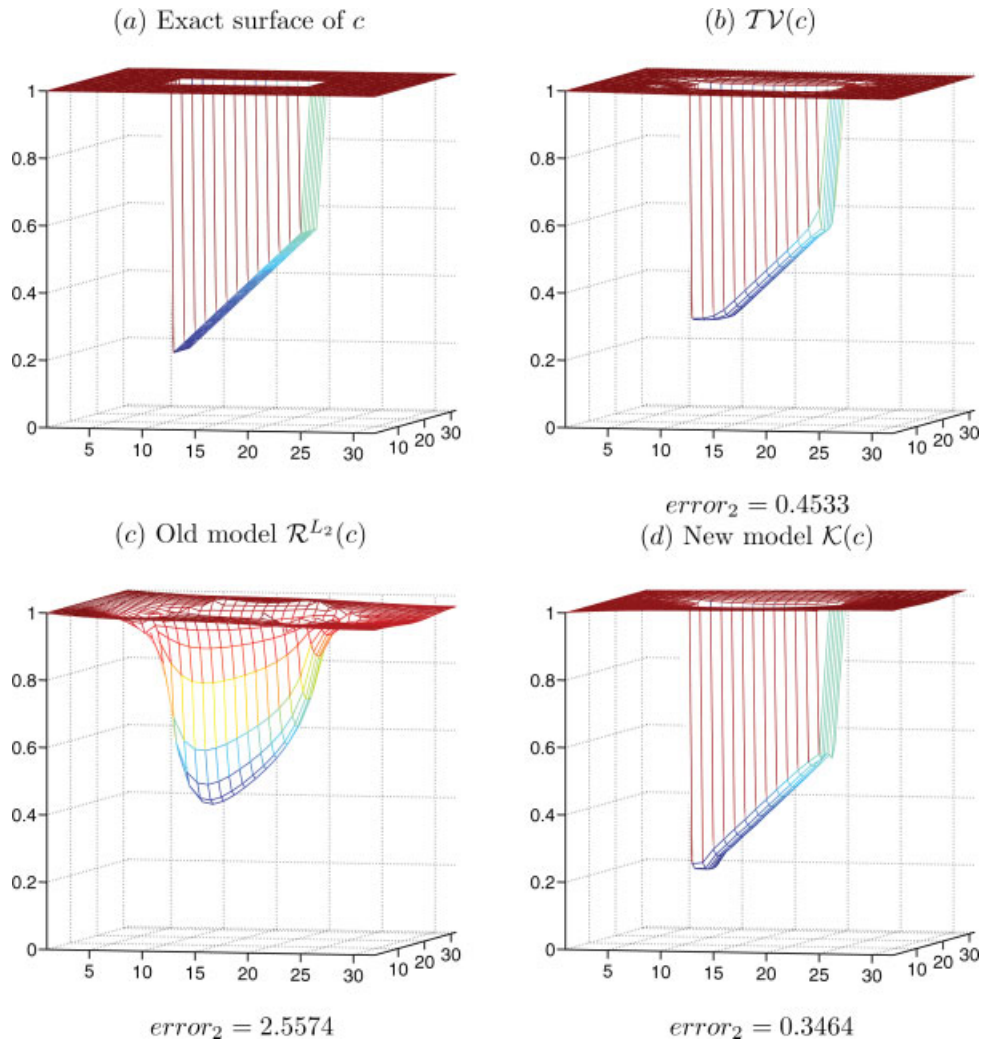

FIG. 6. Surface plots of c for the registration problem in Fig. 5(a, b). (a) The exact surface of c; (b, d) the results by $\mathcal{TV}(c)$, $\mathcal{R}^{L_2}(c)$, and $\mathcal{K}(c)$, respectively. Here, the error_2 denotes the two-norm of the differences between the exact and approximate solutions. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

and

$$\langle \nabla c, \boldsymbol{n} \rangle_{\mathbb{R}^2} = 0, \qquad \langle \nabla \Psi'(\kappa(c)), \boldsymbol{n} \rangle_{\mathbb{R}^2} = 0 \quad \text{on } \partial\Omega. \tag{21}$$

Recall that the first and second terms in (19) are related to the first variations of $\mathcal{D}$ and $\mathcal{R}_l$, respectively.

As used in Ref. [38], the PD idea is suitable for solving a system of higher order nonlinear PDEs like (19). The main idea is to reduce the order and nonlinearity of (19) using the new dual

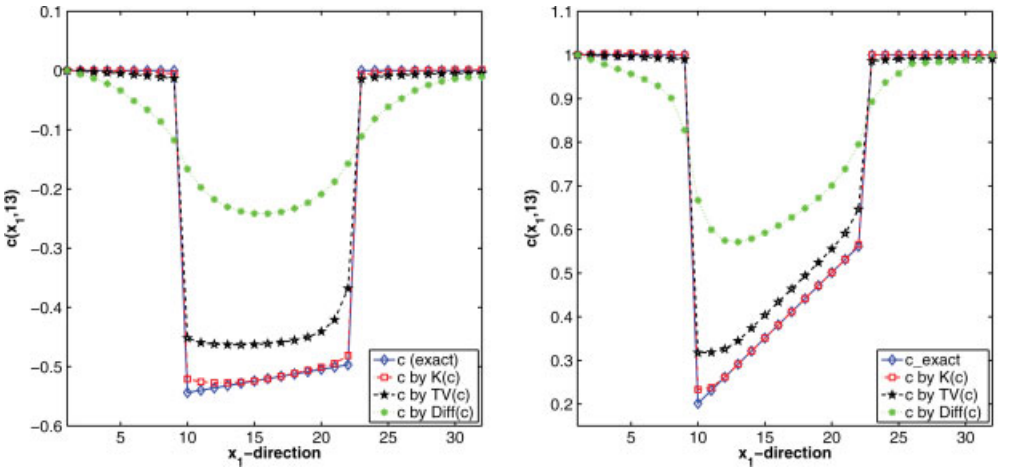

FIG. 7. Plots of the 13th row of $c(x_1, x_2)$ in Figs. 4(a–d) and 6(a–d) by $\mathcal{TV}(c)$, $\mathcal{R}^{L_2}(c)$, and $\mathcal{K}(c)$, respectively. Clearly, our technique $\mathcal{K}(c)$ is an improvement over $\mathcal{TV}(c)$ and $\mathcal{R}^{L_2}(c)$. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

variables. Introducing additional unknown variables (dual variables)

$$v_1 = -\Phi'(\kappa(u_1)) = -\nabla \cdot \frac{\nabla u_1}{|\nabla u_1|_{\beta_1}}, \qquad v_2 = -\Phi'(\kappa(u_2)) = -\nabla \cdot \frac{\nabla u_2}{|\nabla u_2|_{\beta_1}},$$

$$v_3 = -\Psi'(\kappa(c)) = -\nabla \cdot \frac{\nabla c}{|\nabla c|_{\beta_2}}$$

leads (19) to the equivalent system of six second-order nonlinear PDEs given by

$$\begin{cases} -\nabla \cdot \dfrac{\nabla u_1}{|\nabla u_1|_{\beta_1}} - v_1 = g_1 \\ -\nabla \cdot \dfrac{\nabla u_2}{|\nabla u_2|_{\beta_1}} - v_2 = g_2 \\ -\nabla \cdot \dfrac{\nabla c}{|\nabla c|_{\beta_2}} - v_3 = g_3 \\ f_1(\boldsymbol{u}, c) - \alpha_1 \nabla \cdot \left(\dfrac{\nabla v_1}{|\nabla u_1|_{\beta_1}} + \dfrac{\nabla u_1 \cdot (-\nabla v_{1)}}{|\nabla u_1|^3_{\beta_1}} \nabla u_1 \right) = g_4 \\ f_2(\boldsymbol{u}, c) - \alpha_1 \nabla \cdot \left(\dfrac{\nabla v_2}{|\nabla u_2|_{\beta_1}} + \dfrac{\nabla u_2 \cdot (-\nabla v_2)}{|\nabla u_2|^3_{\beta_1}} \nabla u_2 \right) = g_5 \\ f_3(\boldsymbol{u}, c) - \alpha_2 \nabla \cdot \left(\dfrac{\nabla v_3}{|\nabla c|_{\beta_2}} + \dfrac{\nabla c \cdot (-\nabla v_3)}{|\nabla c|^3_{\beta_2}} \nabla c \right) = g_6 \end{cases} \tag{22}$$

subject to the boundary conditions transferred into

$$\langle \nabla u_1, \boldsymbol{n} \rangle_{\mathbb{R}^2} = \langle \nabla u_2, \boldsymbol{n} \rangle_{\mathbb{R}^2} = \langle \nabla c, \boldsymbol{n} \rangle_{\mathbb{R}^2} = \langle \nabla v_m, \boldsymbol{n} \rangle = 0 \quad \text{for } m = 1, 2, 3$$

where $\boldsymbol{g} = (g_1, g_2, g_3, g_4, g_5, g_6)^\top = \boldsymbol{0}$ on the finest grid for the MG setting in the coming section.

Next, we consider our second case in (14), that is, $\mathcal{D} = \mathcal{D}_{\mathrm{II}}^{\mathrm{SSD}}$, $\mathcal{R}_1(\boldsymbol{u}) = \mathcal{R}^{\mathrm{NewCv}}(\boldsymbol{u})$ and $\mathcal{R}_2(c) = \mathcal{K}(c)$ or

$$\min_{\boldsymbol{u},c}\left\{\mathcal{J}_{\alpha_1,\alpha_2}(\boldsymbol{u},c) = \frac{1}{2}\int_\Omega (c(\mathbf{x})T(\mathbf{x}+\boldsymbol{u}(\mathbf{x})) - R(\mathbf{x}))^2 d\mathbf{x} + \alpha_1 \sum_{l=1}^{2}\int_\Omega \Phi(\kappa(u_l))d\mathbf{x} + \alpha_2 \int_\Omega \Phi(\kappa(c))d\mathbf{x}\right\}.$$

Then, it is not difficult to see that similar equations are derived. In fact, only $f_l(\boldsymbol{u},c)$ $(l = 1,2,3)$ in (19) and (22) need modifying and they are to be substituted respectively by

$$f_1(\boldsymbol{u},c) = c(cT(\boldsymbol{u}) - R)\partial_{u_1}T(\boldsymbol{u}), \tag{23}$$

$$f_2(\boldsymbol{u},c) = c(cT(\boldsymbol{u}) - R)\partial_{u_2}T(\boldsymbol{u}), \tag{24}$$

and

$$f_3(\boldsymbol{u},c) = (cT(\boldsymbol{u}) - R)T(\boldsymbol{u}). \tag{25}$$

Refer to Ref. [13] for a discussion of the SSD fitting term.

In this work, we shall mainly consider the first new model (19) with $\mathcal{D} = \mathcal{D}_{\mathrm{I}}^{\mathrm{SSD}}$ in our numerical implementations because it is simpler than the second new model $\mathcal{D} = \mathcal{D}_{\mathrm{II}}^{\mathrm{SSD}}$.

Finally, we remark that the adoption of curvature regularization substantially complicates the resulting PDEs. To illustrate this point, later, we consider replacing the second regularizer $\mathcal{R}_2(c) = \mathcal{K}(c)$ in (18) by

i. $\mathcal{R}_2(c) = \mathcal{TV}(c)$. Then, the last equation in (19) and the boundary conditions in (21) are replaced respectively by

$$f_3(\boldsymbol{u},c) - \alpha_2\nabla\cdot\frac{\nabla c}{|\nabla c|_{\beta_2}} = 0 \tag{26}$$

and $\langle\nabla c, \boldsymbol{n}\rangle_{\mathbb{R}^2} = 0$ on $\partial\Omega$. The resulting PD formulation is then given by

$$\begin{cases} -\nabla\cdot\dfrac{\nabla u_1}{|\nabla u_1|_{\beta_1}} - v_1 = g_1 \\ -\nabla\cdot\dfrac{\nabla u_2}{|\nabla u_2|_{\beta_1}} - v_2 = g_2 \\ f_1(\boldsymbol{u},c) - \alpha_1\nabla\cdot\left(\dfrac{\nabla v_1}{|\nabla u_1|_{\beta_1}} + \dfrac{\nabla u_1\cdot(-\nabla v_1)}{|\nabla u_1|_{\beta_1}^3}\nabla u_1\right) = g_3. \\ f_2(\boldsymbol{u},c) - \alpha_1\nabla\cdot\left(\dfrac{\nabla v_2}{|\nabla u_2|_{\beta_1}} + \dfrac{\nabla u_2\cdot(-\nabla v_2)}{|\nabla u_2|_{\beta_1}^3}\nabla u_2\right) = g_4 \\ f_3(\boldsymbol{u},c) - \alpha_2\nabla\cdot\dfrac{\nabla c}{|\nabla c|_{\beta_2}} = g_5 \end{cases} \tag{27}$$

ii. $\mathcal{R}_2(c) = \mathcal{R}^{L_2}(c)$. Then, the last equation in (19) and the boundary conditions in (21) become

$$f_3(\boldsymbol{u},c) - \alpha_2\Delta c = 0 \tag{28}$$

and $\langle \nabla c, \boldsymbol{n} \rangle_{\mathbb{R}^2} = 0$ on $\partial\Omega$, respectively. Similarly, the resulting PD formulation is given by replacing the last equation in (27) by an easier equation $f_3(\boldsymbol{u}, c) - \alpha_2 \Delta c = g_5$.

Of course, simpler PDEs lead to an easy choice of solvers but do not lead to good registration results.

IV. NUMERICAL SOLUTION FOR THE FORMULATION

To obtain a fast numerical solution of the new formulation (22), we first discuss its discretization and second propose a so-called PDFP iterative method. The unilevel method is further used to develop our MG algorithm. For other MG work in solving various variational imaging models, refer to Refs. [20–22, 39–44].

A. Finite Difference Discretization

For sake of clarity, let $(z_{\widehat{l}}^h)_{i,j} = z_{\widehat{l}}^h(x_{1_i}, x_{2_j})$ denote the grid functions for $\widehat{l} = 1, \ldots, 6$ where

$$\boldsymbol{z} = (z_1, z_2, z_3, z_4, z_5, z_6)^\top = (u_1, u_2, c, v_1, v_2, v_3)^\top \tag{29}$$

and let

$$\Omega_h = \{\mathbf{x} \in \Omega | \mathbf{x} = (x_{1_i}, x_{2_j})^\top = ((2i-1)h/2, (2j-1)h/2), 1 \le i, j \le n\} \tag{30}$$

be the discrete domain consisting of $N = n^2$ cells of size $h \times h$ with the grid mesh $h = 1/n$. The cell-centered finite difference approximations are used with the divergence terms $\nabla \cdot \boldsymbol{V}$ for any vector $\boldsymbol{V} = (V_1, V_2)$ in (22) at a grid point (i, j) as follows:

$$\left(\frac{\partial V_1}{\partial x_1}\right)_{i,j} + \left(\frac{\partial V_2}{\partial x_2}\right)_{i,j} = \frac{(V_1)_{i+1,j} - (V_1)_{i,j}}{h} + \frac{(V_2)_{i,j+1} - (V_2)_{i,j}}{h}. \tag{31}$$

Therefore, we need to calculate V_1 at the grid points $(i+1, j)$ and (i, j) and V_2 at the grid points $(i, j+1)$ and (i, j). We list here the approximations used in our numerical implementations for estimating V_1 at the grid point (i, j) as the following [discretization for V_1 at the grid point $(i+1, j)$ and V_2 at the grid points $(i, j+1)$ and (i, j) can be given similarly]:

$$\kappa\left(z_{\widehat{l}}^h\right)_{i,j} = \left(\nabla \cdot \frac{\nabla z_{\widehat{l}}^h}{\left|\nabla z_{\widehat{l}}^h\right|_{\beta_*}}\right)_{i,j}$$

$$\left(\nabla \cdot \frac{\nabla z_{\widehat{l}}^h}{\left|\nabla z_{\widehat{l}}^h\right|_{\beta_*}}\right)_{i,j} = \left[\frac{\delta_{x_1}^-}{h}\left(\frac{\delta_{x_1}^+\left(z_{\widehat{l}}^h\right)_{i,j}/h}{\sqrt{\beta_* + \left(\delta_{x_1}^+\left(z_{\widehat{l}}^h\right)_{i,j}/h\right)^2 + \left(\delta_{x_2}^+\left(z_{\widehat{l}}^h\right)_{i,j}/h\right)^2}}\right)\right.$$

$$\left. + \frac{\delta_{x_2}^-}{h}\left(\frac{\delta_{x_2}^+\left(z_{\widehat{l}}^h\right)_{i,j}/h}{\sqrt{\beta_* + \left(\delta_{x_1}^+\left(z_{\widehat{l}}^h\right)_{i,j}/h\right)^2 + \left(\delta_{x_2}^+\left(z_{\widehat{l}}^h\right)_{i,j}/h\right)^2}}\right)\right],$$

$$= (1/h^2)\left(\left(\Sigma_{\widehat{l}}^h\right)_{i,j}\left(z_{\widehat{l}}^h\right)_{i,j} - \left(\overline{\Sigma}_{\widehat{l}}^h\right)_{i,j}\left(z_{\widehat{l}}^h\right)_{i,j}\right),$$

$$\left(\Sigma_{\widehat{l}}^h\right)_{i,j} = 2D_{\widehat{l}3}\left(z_{\widehat{l}}^h\right)_{i,j} + D_{\widehat{l}1}\left(z_{\widehat{l}}^h\right)_{i-1,j} + D_{\widehat{l}2}\left(z_{\widehat{l}}^h\right)_{i,j-1},$$

$$
\begin{aligned}
\big(\overline{\Sigma}_{\widehat{l}}^{h}\big)_{i,j}\big(z_{\widehat{l}}^{h}\big)_{i,j} &= \big(D_{\widehat{l}3}\big(z_{\widehat{l}}^{h}\big)_{i,j}\big(\big(z_{\widehat{l}}^{h}\big)_{i+1,j}+\big(z_{\widehat{l}}^{h}\big)_{i,j+1}\big)+D_{\widehat{l}1}\big(z_{\widehat{l}}^{h}\big)_{i,j}\big(z_{\widehat{l}}^{h}\big)_{i-1,j}+D_{\widehat{l}2}\big(z_{\widehat{l}}^{h}\big)_{i,j}\big(z_{\widehat{l}}^{h}\big)_{i,j-1}\big)\big),\\
D_{\widehat{l}1}\big(z_{\widehat{l}}^{h}\big)_{i,j} &= D_{\widehat{l}}\big(z_{\widehat{l}}^{h}\big)_{i-1,j}, \qquad D_{\widehat{l}2}\big(z_{\widehat{l}}^{h}\big)_{i,j}=D_{\widehat{l}}\big(z_{\widehat{l}}^{h}\big)_{i,j-1},\\
D_{\widehat{l}3}\big(z_{\widehat{l}}^{h}\big)_{i,j} &= D_{\widehat{l}}\big(z_{\widehat{l}}^{h}\big)_{i+1,j}=D_{\widehat{l}}\big(z_{\widehat{l}}^{h}\big)_{i,j+1}=D_{\widehat{l}}\big(z_{\widehat{l}}^{h}\big)_{i,j},\\
D_{\widehat{l}}\big(z_{\widehat{l}}^{h}\big)_{i,j} &= \big|\nabla\big(z_{\widehat{l}}^{h}\big)_{i,j}\big|_{\beta_*},\\
\big|\nabla\big(z_{\widehat{l}}^{h}\big)_{i,j}\big|_{\beta_*} &= \sqrt{\beta_*+\big(\delta_{x_1}^{+}\big(z_{\widehat{l}}^{h}\big)_{i,j}/h\big)^2+\big(\delta_{x_2}^{+}\big(z_{\widehat{l}}^{h}\big)_{i,j}/h\big)^2}, \quad \beta_*=\beta_1 \text{ or } \beta_2,\\
\big(z_{\widehat{l}_{x_1}}^{h}\big)_{i,j} &= \delta_{x_1}^{+}\big(z_{\widehat{l}}^{h}\big)_{i,j}/h, \qquad \big(z_{\widehat{l}_{x_2}}^{h}\big)_{i,j}=\delta_{x_2}^{+}\big(z_{\widehat{l}}^{h}\big)_{i,j}/h,\\
\delta_{x_1}^{\pm}\big(z_{\widehat{l}}^{h}\big)_{i,j} &= \pm\big(\big(z_{\widehat{l}}^{h}\big)_{i\pm1,j}-\big(z_{\widehat{l}}^{h}\big)_{i,j}\big), \qquad \delta_{x_2}^{\pm}\big(z_{\widehat{l}}^{h}\big)_{i,j}=\pm\big(\big(z_{\widehat{l}}^{h}\big)_{i,j\pm1}-\big(z_{\widehat{l}}^{h}\big)_{i,j}\big),\\
T_{i,j}^{h^*} &= T^h\big(i+\big(u_1^h\big)_{i,j}, j+\big(u_2^h\big)_{i,j}\big),\\
f_1^h\big(u_1^h,u_2^h,c^h\big)_{i,j} &= \big(T_{i,j}^{h^*}+c_{i,j}-R_{i,j}^h\big)\big(\big(T_{i+1,j}^{h^*}-T_{i-1,j}^{h^*}\big)/(2h)\big),\\
f_2^h\big(u_1^h,u_2^h,c^h\big)_{i,j} &= \big(T_{i,j}^{h^*}+c_{i,j}-R_{i,j}^h\big)\big(\big(T_{i,j+1}^{h^*}-T_{i,j-1}^{h^*}\big)/(2h)\big),\\
f_3^h\big(u_1^h,u_2^h,c^h\big)_{i,j} &= \big(T_{i,j}^{h^*}+c_{i,j}-R_{i,j}^h\big).
\end{aligned}
$$

We note that the finite difference approximations for (22) need to be modified at grid points near the image boundary $\partial\Omega_h$ using the homogeneous Neumann boundary conditions approximated by one-side differences for boundary derivatives:

$$
\big(z_{\widehat{l}}^{h}\big)_{i,1}=\big(z_{\widehat{l}}^{h}\big)_{i,2}, \qquad \big(z_{\widehat{l}}^{h}\big)_{i,n}=\big(z_{\widehat{l}}^{h}\big)_{i,n-1}, \qquad \big(z_{\widehat{l}}^{h}\big)_{1,j}=\big(z_{\widehat{l}}^{h}\big)_{2,j}, \qquad \big(z_{\widehat{l}}^{h}\big)_{n,j}=\big(z_{\widehat{l}}^{h}\big)_{n-1,j}. \quad (32)
$$

B. PDFP Iterative Method

We now discuss a numerical scheme PDFP to solve the discrete version of (22). This is done in two steps:

Outer Iteration Step. First, we introduce a new FP or outer iteration to (22). This can be done as follows; our scheme is semiimplicit in both regularization and data terms. The semiimplicit scheme for the regularization terms is iterated by freezing some coefficients in a similar way to the so-called Lagged-diffusivity method [40] or Quasi-Newton scheme [43, 44]. Staring with an initial guess $\boldsymbol{u}^{[0]}$ (e.g. $\boldsymbol{u}^{[0]}=0$) leads to

$$
\mathbf{N}[z^{[\nu]}]z^{[\nu+1]}=\mathbf{G}[z^{[\nu]}] \quad (33)
$$

where the typical Taylor's expansion for $f_l(u_1^{[\nu+1]},u_2^{[\nu+1]},c^{[\nu+1]})$ of type

$$
\begin{aligned}
f_l\big(u_1^{[\nu+1]},u_2^{[\nu+1]},c^{[\nu+1]}\big) &\approx f_l\big(c^{[\nu]},u_1^{[\nu]},u_2^{[\nu]}\big)+\partial_{u_1}f_l\big(u_1^{[\nu]},u_2^{[\nu]},c^{[\nu]}\big)\delta u_1^{[\nu]}\\
&\quad+\partial_{u_2}f_l\big(u_1^{[\nu]},u_2^{[\nu]},c^{[\nu]}\big)\delta u_2^{[\nu]}+\partial_c f_l\big(u_1^{[\nu]},u_2^{[\nu]},c^{[\nu]}\big)\delta c^{[\nu]}\\
&= f_l\big(u_1^{[\nu]},u_2^{[\nu]},c^{[\nu]}\big)+\sigma_{l1}^{[\nu]}\big(u_1^{[\nu+1]}-u_1^{[\nu]}\big)+\sigma_{l2}^{[\nu]}\big(u_2^{[\nu+1]}-u_2^{[\nu]}\big)\\
&\quad+\sigma_{l3}^{[\nu]}\big(c^{[\nu+1]}-c^{[\nu]}\big) \quad (34)
\end{aligned}
$$

is used in the global linearization scheme. Here,

$$\begin{aligned}\sigma_{l1}^{[\nu]} &= \partial_{u_1} f_l\big(c^{[\nu]}, u_1^{[\nu]}, u_2^{[\nu]}\big)\\ &= (\partial_{u_l} T(\boldsymbol{u}^{[\nu]}))(\partial_{u_1} T(\boldsymbol{u}^{[\nu]})) + (T(\boldsymbol{u}^{[\nu]}) + c^{[\nu]} - R)(\partial_{u_1 u_l} T(\boldsymbol{u}^{[\nu]})), \qquad (35)\\ \sigma_{l2}^{[\nu]} &= \partial_{u_2} f_l\big(c^{[\nu]}, u_1^{[\nu]}, u_2^{[\nu]}\big)\\ &= (\partial_{u_l} T(\boldsymbol{u}^{[\nu]}))(\partial_{u_2} T(\boldsymbol{u}^{[\nu]})) + (T(\boldsymbol{u}^{[\nu]}) + c^{[\nu]} - R)(\partial_{u_2 u_l} T(\boldsymbol{u}^{[\nu]})), \qquad (36)\\ \sigma_{l3}^{[\nu]} &= \partial_c f_l\big(c^{[\nu]}, u_1^{[\nu]}, u_2^{[\nu]}\big) = (\partial_{u_l} T(\boldsymbol{u}^{[\nu]})), \qquad (37)\end{aligned}$$

for $l = 1, 2$ and

$$\sigma_{31}^{[\nu]} = \partial_{u_1} f_3\big(c^{[\nu]}, u_1^{[\nu]}, u_2^{[\nu]}\big) = \partial_{u_1} T(\boldsymbol{u}^{[\nu]}), \qquad (38)$$

$$\sigma_{32}^{[\nu]} = \partial_{u_2} f_3\big(c^{[\nu]}, u_1^{[\nu]}, u_2^{[\nu]}\big) = \partial_{u_2} T(\boldsymbol{u}^{[\nu]}), \qquad (39)$$

$$\sigma_{33}^{[\nu]} = \partial_c f_3\big(c^{[\nu]}, u_1^{[\nu]}, u_2^{[\nu]}\big) = 1. \qquad (40)$$

$$\mathbf{N}[z^{[\nu]}] = \begin{bmatrix} -\mathcal{L}_1\big[u_1^{[\nu]}\big] & 0 & 0 & -1 & 0 & 0\\ 0 & -\mathcal{L}_2\big[u_2^{[\nu]}\big] & 0 & 0 & -1 & 0\\ 0 & 0 & -\mathcal{L}_3[c^{[\nu]}] & 0 & 0 & -1\\ \sigma_{11}^{[\nu]} & \sigma_{12}^{[\nu]} & \sigma_{13}^{[\nu]} & -\alpha_1\mathcal{L}_1\big[u_1^{[\nu]}\big] & 0 & 0\\ \sigma_{21}^{[\nu]} & \sigma_{22}^{[\nu]} & \sigma_{23}^{[\nu]} & 0 & -\alpha_1\mathcal{L}_2\big[u_2^{[\nu]}\big] & 0\\ \sigma_{31}^{[\nu]} & \sigma_{32}^{[\nu]} & \sigma_{33}^{[\nu]} & 0 & 0 & -\alpha_2\mathcal{L}_3[c^{[\nu]}] \end{bmatrix} \qquad (41)$$

$$\begin{aligned} z^{[\nu+1]} &= \big(z_1^{[\nu+1]}, z_2^{[\nu+1]}, z_3^{[\nu+1]}, z_4^{[\nu+1]}, z_5^{[\nu+1]}, z_6^{[\nu+1]}\big)^\top,\\ &= \big(u_1^{[\nu+1]}, u_2^{[\nu+1]}, c^{[\nu+1]}, v_1^{[\nu+1]}, v_2^{[\nu+1]}, v_3^{[\nu+1]}\big)^\top, \qquad (42)\\ \mathbf{G}[z^{[\nu]}] &= \boldsymbol{g} = \big(g_1, g_2, g_3, \widehat{g}_4^{[\nu]}, \widehat{g}_5^{[\nu]}, \widehat{g}_6^{[\nu]}\big)^\top, \qquad (43)\end{aligned}$$

$$\widehat{g}_4^{[\nu]} = g_4 - f_1\big(u_1^{[\nu]}, u_2^{[\nu]}, c^{[\nu]}\big) + \sigma_{11}^{[\nu]} u_1^{[\nu]} + \sigma_{12}^{[\nu]} u_2^{[\nu]} + \sigma_{13}^{[\nu]} c^{[\nu]} + \alpha_1 \nabla \cdot \left(\frac{\nabla u_1^{[\nu]} \cdot (-\nabla v_1)^{[\nu]}}{\big|\nabla u_1^{[\nu]}\big|_\beta^3} \nabla u_1^{[\nu]} \right), \qquad (44)$$

$$\widehat{g}_5^{[\nu]} = g_5 - f_2\big(u_1^{[\nu]}, u_2^{[\nu]}, c^{[\nu]}\big) + \sigma_{21}^{[\nu]} u_1^{[\nu]} + \sigma_{22}^{[\nu]} u_2^{[\nu]} + \sigma_{23}^{[\nu]} c^{[\nu]} + \alpha_1 \nabla \cdot \left(\frac{\nabla u_2^{[\nu]} \cdot (-\nabla v_2)^{[\nu]}}{\big|\nabla u_2^{[\nu]}\big|_\beta^3} \nabla u_2^{[\nu]} \right), \qquad (45)$$

$$\widehat{g}_6^{[\nu]} = g_6 - f_3\big(u_1^{[\nu]}, u_2^{[\nu]}, c^{[\nu]}\big) + \sigma_{31}^{[\nu]} u_1^{[\nu]} + \sigma_{32}^{[\nu]} u_2^{[\nu]} + \sigma_{33}^{[\nu]} c^{[\nu]} + \alpha_2 \nabla \cdot \left(\frac{\nabla c^{[\nu]} \cdot (-\nabla v_3)^{[\nu]}}{|\nabla c^{[\nu]}|_\beta^3} \nabla c^{[\nu]} \right), \qquad (46)$$

and

$$\mathcal{L}_m\big[z_m^{[\nu]}\big]z_{\widehat{l}}^{[\nu+1]} = \nabla\cdot\left(\frac{\overbrace{D_m\big(z_m^{[\nu]}\big)}}{\big|\nabla z_m^{[\nu]}\big|_{\beta_1}}\nabla z_{\widehat{l}}^{[\nu+1]}\right) \quad (m=1,2,3 \text{ and } \widehat{l}=m \text{ or } \widehat{l}=m+3). \tag{47}$$

Inner Iteration Step. After applying the finite difference approximations represented in Subsection A of section 4 with (33), the so-called pointwise collective Gauss–Seidel (PCGS) relaxation method is used as the inner solver to solve inexactly the associated linear system. Here, the kth PCGS step is given by

$$(z^{[\nu+1]})_{i,j}^{[k+1]} = (\mathbf{N}[z^{[\nu]}]_{i,j})^{-1}(\mathbf{G}[z^{[\nu]}])_{i,j}^{[k+1/2]}, \tag{48}$$

where the symbol of the mesh parameter h is dropped for simplicity. We note that other choices of iterative techniques such as the line relaxation techniques or the preconditioned conjugate gradient method are optional. However, they are computationally more expensive than the PCGS relaxation method.

Numerical results have shown that the PCGS relaxation method is not suitable as a potential (MG) smoother, in particular for nonsmooth problems, since high values of the smoothing factors appear, especially around large local variations in the coefficients $D_m(z_m^{[\nu]})_{i,j}$. To avoid this situation, we introduce the so-called relaxation parameter $\omega \in (0,2)$, typically $\omega = 0.7$, and iterate the $\omega-$PCGS steps at those odd points by

$$(z^{[\nu+1]})_{i,j}^{[k+1]} = (1-\omega)(z^{[\nu+1]})_{i,j}^{[k]} + \omega\underbrace{(\mathbf{N}[z^{[\nu]}]_{i,j})^{-1}(\mathbf{G}[z^{[\nu]}])_{i,j}^{[k+1/2]}}_{\text{original PCGS result}}, \tag{49}$$

with the following notation

$$\mathbf{N}[z^{[\nu]}]_{i,j} = \frac{1}{h^2}\begin{bmatrix} \big(\Sigma_1^{[\nu]}\big)_{i,j} & 0 & 0 & -1 & 0 & 0 \\ 0 & \big(\Sigma_2^{[\nu]}\big)_{i,j} & 0 & 0 & -1 & 0 \\ 0 & 0 & \big(\Sigma_3^{[\nu]}\big)_{i,j} & 0 & 0 & -1 \\ h^2\sigma_{11}^{[\nu]} & h^2\sigma_{12}^{[\nu]} & h^2\sigma_{13}^{[\nu]} & \alpha_1\big(\Sigma_1^{[\nu]}\big)_{i,j} & 0 & 0 \\ h^2\sigma_{21}^{[\nu]} & h^2\sigma_{22}^{[\nu]} & h^2\sigma_{23}^{[\nu]} & 0 & \alpha_1(\Sigma_2^{[\nu]})_{i,j} & 0 \\ h^2\sigma_{31}^{[\nu]} & h^2\sigma_{32}^{[\nu]} & h^2\sigma_{33}^{[\nu]} & 0 & 0 & \alpha_2\big(\Sigma_3^{[\nu]}\big)_{i,j} \end{bmatrix} \tag{50}$$

and

$$(\mathbf{G}[z^{[\nu]}])_{i,j}^{[k+1/2]} = \begin{pmatrix} (g_1)_{i,j} + (1/h^2)\big(\overline{\Sigma}_1^{[\nu]}\big)_{i,j}\big(u_1^{[\nu+1]}\big)_{i,j}^{[k+1/2]} \\ (g_2)_{i,j} + (1/h^2)\big(\overline{\Sigma}_2^{[\nu]}\big)_{i,j}\big(u_2^{[\nu+1]}\big)_{i,j}^{[k+1/2]} \\ (g_3)_{i,j} + (1/h^2)\big(\overline{\Sigma}_3^{[\nu]}\big)_{i,j}(c^{[\nu+1]})_{i,j}^{[k+1/2]} \\ (\widehat{g}_4)_{i,j}^{[\nu]} + (\alpha_1/h^2)\big(\overline{\Sigma}_1^{[\nu]}\big)_{i,j}\big(v_1^{[\nu+1]}\big)_{i,j}^{[k+1/2]} \\ (\widehat{g}_5)_{i,j}^{[\nu]} + (\alpha_1/h^2)\big(\overline{\Sigma}_2^{[\nu]}\big)_{i,j}\big(v_2^{[\nu+1]}\big)_{i,j}^{[k+1/2]} \\ (\widehat{g}_6)_{i,j}^{[\nu]} + (\alpha_2/h^2)\big(\overline{\Sigma}_3^{[\nu]}\big)_{i,j}\big(v_3^{[\nu+1]}\big)_{i,j}^{[k+1/2]} \end{pmatrix}. \tag{51}$$

Finally, our proposed solver can be summarized as follows:

Algorithm 1. (our proposed iterative solver: PDFP)

Define:
A regularization parameter, α
A relaxation parameter, ω
$K > 0$ tolerance (typically $K = 0.5/\sqrt{\beta^*}$ where $\beta^* = \min\{\beta_1, \beta_2\}$)
PCGSiter the maximum number of PCGS iterations

$$[\overline{z}^h] \leftarrow \text{Solver}(\overline{z}^h, \boldsymbol{g}^h, R^h, T^h, \alpha, \omega, K, \text{PCGSiter})$$

- Use input parameters to compute $(\sigma_{lm})_{i,j}$, $(\mathbf{G}_h[\overline{z}^h])_{i,j}$, and $(\mathbf{N}_h[\overline{z}^h]_{i,j})^{-1}$ for $l, m = 1, 2, 3$ and $1 \le i, j \le n$
- Perform PCGS steps
  - ▶ for $k = 1 : \text{PCGSiter}$
    - ▶ for $i = 1 : n$
      - ▶ for $j = 1 : n$
        - ▶ if $D_m(\overline{z}_m^h)_{i,j} \ge K \cdot \min\{D_{m1}(\overline{z}_m)_{i,j}, D_{m2}(\overline{z}_m)_{i,j}, D_{m3}(\overline{z}_m)_{i,j}\}$ for $m = 1, 2$ or 3
          - ♦ Set $\omega = 0.7$
        - else
          - ♦ Set $\omega = 1.0$
        - end
        - ▶ Compute $(\widetilde{z}^h)_{i,j}^{[k+1]}$ using (48)
        - ▶ $(\overline{z})_{i,j}^{[k+1]} = (1 - \omega)(\overline{z})_{i,j}^{[k]} + \omega(\widetilde{z}^h)_{i,j}^{[k+1]}$
      - ▶ end
    - ▶ end
  - ▶ end

C. NMG Algorithm

MG techniques [45–49] have been proved to be very useful in the context of deformable image registration for solving large systems of linear or nonlinear equations arising from high-resolution digital images in real-life applications. The basic idea of a MG method is to smooth high-frequency components of the error of the solution on a fine grid by performing a few steps with a smoother (an iterative relaxation technique) such that a smooth error term can be well represented and approximated on a coarser grid. After a residual equation has been solved on the coarse grid, a coarse-grid correction is interpolated back to the fine grid and used to correct the fine grid approximation. Finally, the smoother is performed again to remove any new high-frequency components of the error introduced by the interpolation. This is known as a two-grid cycle, and with recursive application it can be extended to a MG method. The NMG method based on Brandt's full approximation scheme (FAS-NMG) is widely used for solving nonlinear problems, in particular image processing applications. For general PDEs, convergence of such a FAS-NMG method is not assured; often a suitable smoother is hard to construct. Fortunately for our new model (19), the above PDFP iterative method turns out to be an effective smoother.

Here, we have to solve (22), a coupled system of six nonlinear PDEs denoted by

$$\mathcal{N}^h(\boldsymbol{z}^h) = \boldsymbol{g}^h, \text{i.e.} \begin{cases} \mathcal{N}_1^h(\boldsymbol{z}^h) = g_1^h \\ \vdots \\ \mathcal{N}_6^h(\boldsymbol{z}^h) = g_6^h \end{cases} \tag{52}$$

and involving the nonlinear partial differential operator $\mathcal{N}_{\widehat{l}}^h(\boldsymbol{u}^h)$ given by the left-hand side of (22), where $g_{\widehat{l}} = 0$ on the finest grid, for $\widehat{l} = 1, \ldots, 6$.

Let $\overline{\boldsymbol{z}}^h = (\overline{z}_1^h, \overline{z}_2^h, \overline{z}_3^h, \overline{z}_4^h, \overline{z}_5^h, \overline{z}_6^h)^\top$ be the approximation of $\boldsymbol{z}^h$ after a few smoothing iterations in a presmoothing step on a fine-grid problem where we denote by $\boldsymbol{z}^h$ the exact solution of (22). Then, the algebraic error $\boldsymbol{e}^h$ of the solution is given by $\boldsymbol{e}^h = \boldsymbol{z}^h - \overline{\boldsymbol{z}}^h$. The residual equation system for the $\widehat{l}^{\text{th}}$ equation is given by

$$\mathcal{N}_{\widehat{l}}^h(\overline{\boldsymbol{z}}^h + \boldsymbol{e}^h) - \mathcal{N}_{\widehat{l}}^h(\overline{\boldsymbol{z}}^h) = g_{\widehat{l}}^h - \mathcal{N}_{\widehat{l}}^h(\overline{\boldsymbol{z}}^h) = r_{\widehat{l}}^h.$$

To correct the approximated solution $\overline{\boldsymbol{z}}^h$ on the fine grid, one needs to compute the error $\boldsymbol{e}^h$. However, the computation of $\boldsymbol{e}^h$ is prohibitively expensive and cannot be computed directly on the fine grid. As high-frequency components of the error in the presmoothing step have already been removed by the smoother, we can transfer the following nonlinear system to the coarse grid as follows:

$$\underbrace{\mathcal{N}_{\widehat{l}}^h(\overline{\boldsymbol{z}}^h + \boldsymbol{e}^h)}_{\mathcal{N}_{\widehat{l}}^h(z^h)} = \underbrace{r_{\widehat{l}}^h + \mathcal{N}_{\widehat{l}}^h(\overline{\boldsymbol{z}}^h)}_{g_{\widehat{l}}^h} \rightarrow \underbrace{\mathcal{N}_{\widehat{l}}^H(\overline{\boldsymbol{z}}^H + \boldsymbol{e}^H)}_{\mathcal{N}_{\widehat{l}}^H(z^H)} = \underbrace{r_{\widehat{l}}^H + \mathcal{N}_{\widehat{l}}^H(\overline{\boldsymbol{z}}^H)}_{g_{\widehat{l}}^H} \tag{53}$$

where $H = 2h$ is the new cell size $H \times H$ and $g_{\widehat{l}}^H \neq 0$ on the coarse grid. After the nonlinear residual equation on the coarse grid (53) has been solved with a method of our choice, the coarse-grid correction $\boldsymbol{e}^H = \boldsymbol{z}^H - \overline{\boldsymbol{z}}^H$ is then interpolated back to the fine grid $\boldsymbol{e}^h$ that can now be used for updating the approximated solution $\overline{\boldsymbol{z}}^h$ of the original system on the fine grid $\overline{\boldsymbol{z}}_{new}^h = \overline{\boldsymbol{z}}^h + \boldsymbol{e}^h$ (coarse-grid correction step). The last step for a FAS-NMG method is to perform the smoother again to remove high-frequency parts of the interpolated error (post-smoothing step).

We now define our MG components as follows. The PDFP method represented in Subsection B of section 4 is applied as the MG smoother. Standard coarsening is used in computing the coarse-grid domain Ω_H by doubling the grid size in each space direction, that is, $h \rightarrow 2h = H$. For intergrid transfer operators between Ω_h and Ω_H, the averaging and bilinear interpolation techniques are used for the restriction and interpolation operators denoted respectively by I_h^H and I_H^h; see the details in Refs. [45–49]. To compute the coarse-grid operator of $\mathcal{N}_{\widehat{l}}^h(\boldsymbol{z}^h)$ given by the left-hand side of (22), a so-called discretization coarse grid approximation is performed [45, 47, 49]. The idea is to rediscretize the Euler–Lagrange system directly.

Finally, the pseudo-code implementation of our FAS-NMG method can be summarized in the following algorithm:

Algorithm 2. (FAS-NMG Algorithm)

Denote the FAS-NMG parameters as follows:

ν_1 presmoothing steps on each level

ν_2 postsmoothing steps on each level

μ the number of MG cycles on each level ($\mu = 1$ for V-cycling and $\mu = 2$ for W-cycling)

[Here, we present the V-cycle with $\mu = 1$]

α a regularization parameter
ω a relaxation parameter
$K > 0$ tolerance
PCGSiter the maximum number of iterations using a smoother

$$\overline{z}^h \leftarrow \mathrm{FASMG}(\overline{z}^h, \alpha, \overrightarrow{\varepsilon})$$

- Select α, $\overrightarrow{\varepsilon} = (\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4)$ and initial guess solutions $\widetilde{z}^h_{\text{initial}}$ on the finest grid
- Set $K = 0$, $(\overline{z}^h)^{[K]} = \widetilde{z}^h_{\text{initial}}$, $\widetilde{\varepsilon}_2 = \varepsilon_2 + 1$, $\widetilde{\varepsilon}_3 = \varepsilon_3 + 1$, and $\widetilde{\varepsilon}_4 = \varepsilon_4 + 1$
- While ($K < \varepsilon_1$ AND $\widetilde{\varepsilon}_2 \geq \varepsilon_2$ AND $\widetilde{\varepsilon}_3 \geq \varepsilon_3$ AND $\widetilde{\varepsilon}_4 \geq \varepsilon_4$)
  - ▶ $(\overline{z}^h)^{[K+1]} \leftarrow \mathrm{FASCYC}((\overline{z}^h)^{[K]}, \boldsymbol{g}^h, R^h, T^h, \nu_1, \nu_2, \alpha, \omega, K, \mathrm{PCGSiter})$
  - ▶ $\widetilde{\varepsilon}_2 = \mathrm{mean}\{\|g^h_{\widehat{l}} - \mathcal{N}^h_{\widehat{l}}((\overline{z}^h)^{[K+1]})\|_2 / \|g^h_{\widehat{l}} - \mathcal{N}^h_{\widehat{l}}(\widetilde{z}^h_{\text{initial}})\|_2 | \widehat{l} = 1, \ldots, 6\}$
  - ▶ $\widetilde{\varepsilon}_3 = \overline{\mathcal{D}}^h(R^h, T^h_\star(\overline{\boldsymbol{u}}^h)^{[K+1]}) / \overline{\mathcal{D}}^h(R^h, T^h_\star(\overline{\boldsymbol{u}}^h)^{[0]})$,
    [Recall that $\overline{\mathcal{D}}^h(R^h, T^h_\star(\cdot)) \sim \frac{h^2}{2}\|R^h, T^h_\star(\cdot)\|_2^2$]
  - ▶ $\widetilde{\varepsilon}_4 = |\overline{\mathcal{D}}^h(R^h, T^h_\star(\overline{\boldsymbol{u}}^h)^{[K+1]}) - \overline{\mathcal{D}}^h(R^h, T^h_\star(\overline{\boldsymbol{u}}^h)^{[K]})|$
  - ▶ $K = K + 1$
- end

---

where

$$[\overline{z}^h] \leftarrow \mathrm{FASCYC}(\overline{z}^h, \boldsymbol{g}^h, R^h, T^h, \nu_1, \nu_2, \alpha, \omega, K, \mathrm{PCGSiter})$$

- If Ω_h = coarset grid ($|\Omega_h| = 8 \times 8$), solve (22) using Algorithm 1 and then stop. Else continue with the following steps.
- Presmoothing:
  For $k = 1$ to ν_1, $[\overline{z}^h] \leftarrow \mathrm{Solver}(\overline{z}^h, \boldsymbol{g}^h, R^h, T^h, \alpha, \omega, K, \mathrm{PCGSiter})$
- Restriction to the coarse grid:
  $\overline{z}^H_{\widehat{l}} \leftarrow I^H_h \overline{z}^h_{\widehat{l}}$ (for $\widehat{l} = 1, \ldots, 6$), $R^H \leftarrow I^H_h R^h, T^H \leftarrow I^H_h T^h$
- Set the initial solution for the coarse-grid problem:
  $\widetilde{z}^H_{\widehat{l}} \leftarrow \overline{z}^H_{\widehat{l}}$
- Compute the new right-hand side for the coarse-grid problem:
  $g^H_{\widehat{l}} \leftarrow I^H_h(g^h_{\widehat{l}} - \mathcal{N}^h_{\widehat{l}}(\overline{z}^h)) + \mathcal{N}^H_{\widehat{l}}(\overline{z}^H)$ (for $\widehat{l} = 1, \ldots, 6$)
- Implement the FAS-NMG method on the coarse-grid problem:
  For $k = 1$ to μ, $[\overline{z}^H] \leftarrow \mathrm{FASCYC}(\overline{z}^H, \boldsymbol{g}^H, R^H, T^H, \nu_1, \nu_2, \alpha, \omega, K, \mathrm{PCGSiter})$
- Add the coarse-grid corrections:
  $\overline{z}^h_{\widehat{l}} \leftarrow \overline{z}^h_{\widehat{l}} + I^h_H(\overline{z}^H_{\widehat{l}} - \widetilde{z}^H_{\widehat{l}})$, (for $\widehat{l} = 1, \ldots, 6$)
- Postsmoothing:
  For $k = 1$ to ν_2, $[\overline{z}^h] \leftarrow \mathrm{Solver}(\overline{z}^h, \boldsymbol{g}^h, R^h, T^h, \alpha, \omega, K, \mathrm{PCGSiter})$

For practical applications our FAS-NMG approach is stopped if the maximum number of V- or W-cycles ε_1 is reached (usually $\varepsilon_1 = 20$), the mean of the relative residuals obtained from the Euler–Lagrange equations (22) is smaller than a small number $\varepsilon_2 > 0$ (typically $\varepsilon_2 = 10^{-4}$), the

relative reduction of the dissimilarity is smaller than some $\varepsilon_3 > 0$ (we usually assign $\varepsilon_3 = 0.10$ meaning that the relative reduction of the dissimilarity would decrease about 90%), or the change in two consecutive steps of the data/fitting term $\mathcal{D}$ is smaller than a small number $\varepsilon_4 > 0$ (typically $\varepsilon_4 = 10^{-4}$).

D. Local Fourier Analysis For The PDFP Method

As mentioned, a FAS-NMG method may not converge for nonlinear PDEs. In fact, we have tried many other approaches for solving (20), to which no effective unilevel iterative methods (as possible smoothers) were found. Only our proposed PDFP method for the reformulated system (22) was found effective. Later we shall analyze the PDFP smoother to find a theoretical reason for its effectiveness using a local Fourier analysis (LFA).

The LFA is a powerful tool for analyzing the smoothing properties of iterative algorithms used in MG methods. Although LFA was originally developed for discrete linear operators with constant coefficients on infinite grids, it can also be applied to more general nonlinear equations with varying coefficients such as the discrete versions of (22). To this end, first an infinite grid is assumed to eliminate the effect of boundary conditions and second it is assumed that the discrete nonlinear operator can be linearized (by freezing coefficients) and replaced locally by a new operator with constant coefficients [47]. This approach has proved to be very useful in the understanding of MG methods when solving nonlinear problems; see for instance Refs. [20–22, 24, 29, 30, 32, 50–52] for interesting examples and discussions.

For linear problems, iterative methods such as damped Jacobi or Gauss–Seidel methods are usually enough to rapidly reduce high frequencies of the underlying error. However, for nonlinear problems, nonstandard smoothers are often required and their efficiency in smoothing is the decisive factor in determining whether a given MG is convergent or not. For nonlinear and anisotropic problems such as (22), developing such an effective smoother is by no means a trivial task. A quantitative measure of the smoothing efficiency for a given algorithm is the smoothing factor, which is defined as the worst asymptotic error reduction, by performing one smoother step, of all high-frequency error components [47, 49], denoted by $\overline{\mu}$ and numerically computed for test problems.

Later we shall use the LFA to analyze the smoothing properties of the PDFP iterations applied to the linearized system $\mathbf{N}_h[\overline{z}^h]z^h = \mathbf{G}_h[\overline{z}^h]$ obtained by freezing coefficients in (33) at some outer step. Here, z^h and $\overline{z}^h$ denote the exact solution and the current approximation and $\mathbf{N}[\overline{z}^h]$ and $\mathbf{G}[\overline{z}^h]$ the resulting discrete operators from the linearization at $\overline{z}^h$.

Let $\boldsymbol{\varphi}_h(\boldsymbol{\theta}, \mathbf{x}) = \exp(\mathbf{i}\boldsymbol{\theta}\mathbf{x}/h) \cdot \widehat{\mathbf{I}}$ be grid functions, where $\widehat{\mathbf{I}} = (1, 1, 1, 1, 1, 1)^\top$, $\boldsymbol{\theta} = (\theta_1, \theta_2)^\top \in \Theta = (-\pi, \pi]^2$, $\mathbf{x} \in \Omega_h^\infty$, and $\mathbf{i} = \sqrt{-1}$. Similarly, our LFA is performed over the infinite grid

$$\Omega_h^\infty = \{\mathbf{x} \in \Omega | \mathbf{x} = (x_{1_i}, x_{2_j})^\top = ((2i-1)h/2, (2j-1)h/2)^\top, i, j \in \mathbb{Z}^2\}. \tag{54}$$

and applied to each grid point $\xi = (i, j)$ separately. Here, we denote by

$$\bar{\mu}_{\text{loc}} = \max_{\xi \in \Omega_h} \mu_{\text{loc}}$$

the smoothing factor defined as the worst possible value of the local smoothing factor $\mu_{\text{loc}} = \mu(\xi)$ over Ω_h and $\mathbf{N}_h(\xi)z^h(\xi) = \mathbf{G}_h(\xi)$ the local discrete system centered and defined only within a small neighborhood of ξ and $\boldsymbol{u}^h(\xi) = [u_1^h(\xi), u_2^h(\xi)]$.

Let us consider first the case of the PCGS ($\omega = 1$) approach. The splitting

$$\mathbf{N}_h(\xi) = \mathbf{N}_h^{[+]}(\xi) + \mathbf{N}_h^{[0]}(\xi) + \mathbf{N}_h^{[-]}(\xi)$$

leads the local inner iterations to

$$\mathbf{N}_h^{[+]}(\xi)\overline{z}_{\text{new}}^h(\xi) + \mathbf{N}_h^{[0]}(\xi)\overline{z}_{\text{new}}^h(\xi) + \mathbf{N}_h^{[-]}(\xi)\overline{z}_{\text{old}}^h(\xi) = \mathbf{G}_h(\xi) \tag{55}$$

where $\overline{z}_{\text{old}}^h(\xi)$ and $\overline{z}_{\text{new}}^h(\xi)$ are the approximations to $z^h(\xi)$ before and after the inner smoothing step, respectively. Here,

$$N_h^{[+]}(\xi) = \begin{bmatrix} -\mathcal{L}_1^{h[+]}(\xi) & 0 & 0 & 0 & 0 & 0 \\ 0 & -\mathcal{L}_2^{h[+]}(\xi) & 0 & 0 & 0 & 0 \\ 0 & 0 & -\mathcal{L}_3^{h[+]}(\xi) & 0 & 0 & 0 \\ 0 & 0 & 0 & -\alpha_1\mathcal{L}_1^{h[+]}(\xi) & 0 & 0 \\ 0 & 0 & 0 & 0 & -\alpha_1\mathcal{L}_2^{h[+]}(\xi) & 0 \\ 0 & 0 & 0 & 0 & 0 & -\alpha_2\mathcal{L}_3^{h[+]}(\xi) \end{bmatrix}, \tag{56}$$

$$\mathbf{N}_h^{[0]}(\xi) = \begin{bmatrix} -\mathcal{L}_1^{h[0]}(\xi) & 0 & 0 & -1 & 0 & 0 \\ 0 & -\mathcal{L}_2^{h[0]}(\xi) & 0 & 0 & -1 & 0 \\ 0 & 0 & -\mathcal{L}_3^{h[0]}(\xi) & 0 & 0 & -1 \\ \sigma_{11}(\xi) & \sigma_{12}(\xi) & \sigma_{13}(\xi) & -\alpha_1\mathcal{L}_1^{h[0]}(\xi) & 0 & 0 \\ \sigma_{21}(\xi) & \sigma_{22}(\xi) & \sigma_{23}(\xi) & 0 & -\alpha_1\mathcal{L}_2^{h[0]}(\xi) & 0 \\ \sigma_{31}(\xi) & \sigma_{32}(\xi) & \sigma_{33}(\xi) & 0 & 0 & -\alpha_2\mathcal{L}_3^{h[0]}(\xi) \end{bmatrix}, \tag{57}$$

$$\mathbf{N}_h^{[-]}(\xi) = \begin{bmatrix} -\mathcal{L}_1^{h[-]}(\xi) & 0 & 0 & 0 & 0 & 0 \\ 0 & -\mathcal{L}_2^{h[-]}(\xi) & 0 & 0 & 0 & 0 \\ 0 & 0 & -\mathcal{L}_3^{h[-]}(\xi) & 0 & 0 & 0 \\ 0 & 0 & 0 & -\alpha_1\mathcal{L}_1^{h[-]}(\xi) & 0 & 0 \\ 0 & 0 & 0 & 0 & -\alpha_1\mathcal{L}_2^{h[-]}(\xi) & 0 \\ 0 & 0 & 0 & 0 & 0 & -\alpha_2\mathcal{L}_3^{h[-]}(\xi) \end{bmatrix}, \tag{58}$$

$$-\mathcal{L}_m^{h[+]}(\xi) = \frac{1}{h^2}\begin{bmatrix} 0 & 0 & 0 \\ -D_{m2}(\overline{z}_m(\xi)) & 0 & 0 \\ 0 & -D_{m1}(\overline{z}_m(\xi)) & 0 \end{bmatrix}, \tag{59}$$

$$-\mathcal{L}_m^{h[0]}(\xi) = \frac{1}{h^2}\begin{bmatrix} 0 & 0 & 0 \\ 0 & \Sigma_m(\xi) & 0 \\ 0 & 0 & 0 \end{bmatrix}, \tag{60}$$

and

$$-\mathcal{L}_m^{h[-]}(\xi) = \frac{1}{h^2}\begin{bmatrix} 0 & -D_{m3}(\overline{z}_l(\xi)) & 0 \\ 0 & 0 & -D_{m3}(\overline{z}_l(\xi)) \\ 0 & 0 & 0 \end{bmatrix}, \tag{61}$$

for $m = 1, 2, 3$. Subtracting (55) from $\mathbf{N}_h(\xi)\mathbf{z}^h(\xi) = \mathbf{G}_h(\xi)$ yields the system of local error equations

$$\mathbf{N}_h^{[+]}(\xi)\overline{\boldsymbol{e}}_{\text{new}}^h(\xi) + \mathbf{N}_h^{[0]}(\xi)\overline{\boldsymbol{e}}_{\text{new}}^h(\xi) + \mathbf{N}_h^{[-]}(\xi)\overline{\boldsymbol{e}}_{\text{old}}^h(\xi) = 0$$

or

$$\overline{\boldsymbol{e}}^h_{\text{new}}(\xi) = \mathbf{S}_h(\xi)\overline{\boldsymbol{e}}^h_{\text{old}}(\xi)$$

where

$$\overline{\boldsymbol{e}}^h_{\text{old}}(\xi) = \boldsymbol{z}^h(\xi) - \overline{\boldsymbol{z}}^h_{\text{old}}(\xi) \text{ and } \overline{\boldsymbol{e}}^h_{\text{new}}(\xi) = \boldsymbol{z}^h(\xi) - \overline{\boldsymbol{z}}^h_{\text{new}}(\xi)$$

are the error functions and

$$\mathbf{S}_h(\xi) = -\left[\mathbf{N}_h^{[0]}(\xi) + \mathbf{N}_h^{[+]}(\xi)\right]^{-1}\left[\mathbf{N}_h^{[-]}(\xi)\right] \tag{62}$$

is the amplification factor. Recall that the Fourier symbols of $\mathcal{L}_l^{h[+]}(\xi)$ and $\mathcal{L}_l^{h[+]}(\xi)$ denoted by

$$-\mathcal{L}_m^{h[+]}(\xi,\boldsymbol{\theta}) = \frac{1}{h^2}(\Sigma_m(\xi) - D_{m1}(\xi)\exp(-\mathbf{i}\theta_1) - D_{m2}(\xi)\exp(-\mathbf{i}\theta_2)) \tag{63}$$

and

$$-\mathcal{L}_m^{h[-]}(\xi,\boldsymbol{\theta}) = -\frac{1}{h^2}(D_{m3}(\xi)(\exp(\mathbf{i}\theta_1) + \exp(\mathbf{i}\theta_2))). \tag{64}$$

are used to compute (65); [47,49]. Hence, the PDFP local smoothing factor for this case is defined by

$$\mu_{\text{loc}} = \sup\{|\rho(\widehat{\mathbf{S}}_h(\xi,\boldsymbol{\theta}))| : \boldsymbol{\theta} \in \Theta_{\text{high}}\} \tag{65}$$

where

$$\widehat{\mathbf{S}}_h(\xi,\boldsymbol{\theta}) = -\left[\widehat{\mathbf{N}}_h^{[0]}(\xi,\boldsymbol{\theta}) + \widehat{\mathbf{N}}_h^{[+]}(\xi,\boldsymbol{\theta})\right]^{-1}\left[\widehat{\mathbf{N}}_h^{[-]}(\xi,\boldsymbol{\theta})\right]$$

is the Fourier symbol of $\mathbf{S}_h(\xi)$.

For the case of the ω-PCGS approach, the PDFP local smoothing factor can be defined in a similar way to (65),

$$\mu_{\text{loc}} = \sup\{|\rho(\widehat{\mathbf{S}}_h(\xi,\boldsymbol{\theta},\omega))| : \boldsymbol{\theta} \in \Theta_{\text{high}}\}, \tag{66}$$

where the Fourier symbol of the amplification factor $\mathbf{S}_h(\xi,\omega)$ is given by

$$\widehat{\mathbf{S}}_h(\xi,\boldsymbol{\theta},\omega) = \left[\widehat{\mathbf{N}}_h^{[0]}(\xi,\boldsymbol{\theta}) + \omega\widehat{\mathbf{N}}_h^{[+]}(\xi,\boldsymbol{\theta})\right]^{-1}\left[(1-\omega)\widehat{\mathbf{N}}_h^{[0]}(\xi,\boldsymbol{\theta}) - \omega\widehat{\mathbf{N}}_h^{[-]}(\xi,\boldsymbol{\theta})\right] \in \mathbb{C}^{6\times 6}. \tag{67}$$

To select the optimal value of ω, we used four registration problems shown in Fig. 8 on a 32×32 grid. Our results indicated that $\omega = 0.7$ provides good smoothing properties ($\overline{\mu}^*_{\text{loc}} \approx 0.60$). Clearly a smoothing rate of 0.6 is definitely sufficient for driving a MG algorithm [47,53]. We also conducted several numerical tests to confirm that (33) is a potential smoother for our FAS-NMG method to solve (22); see Table I in Section B of section 5.

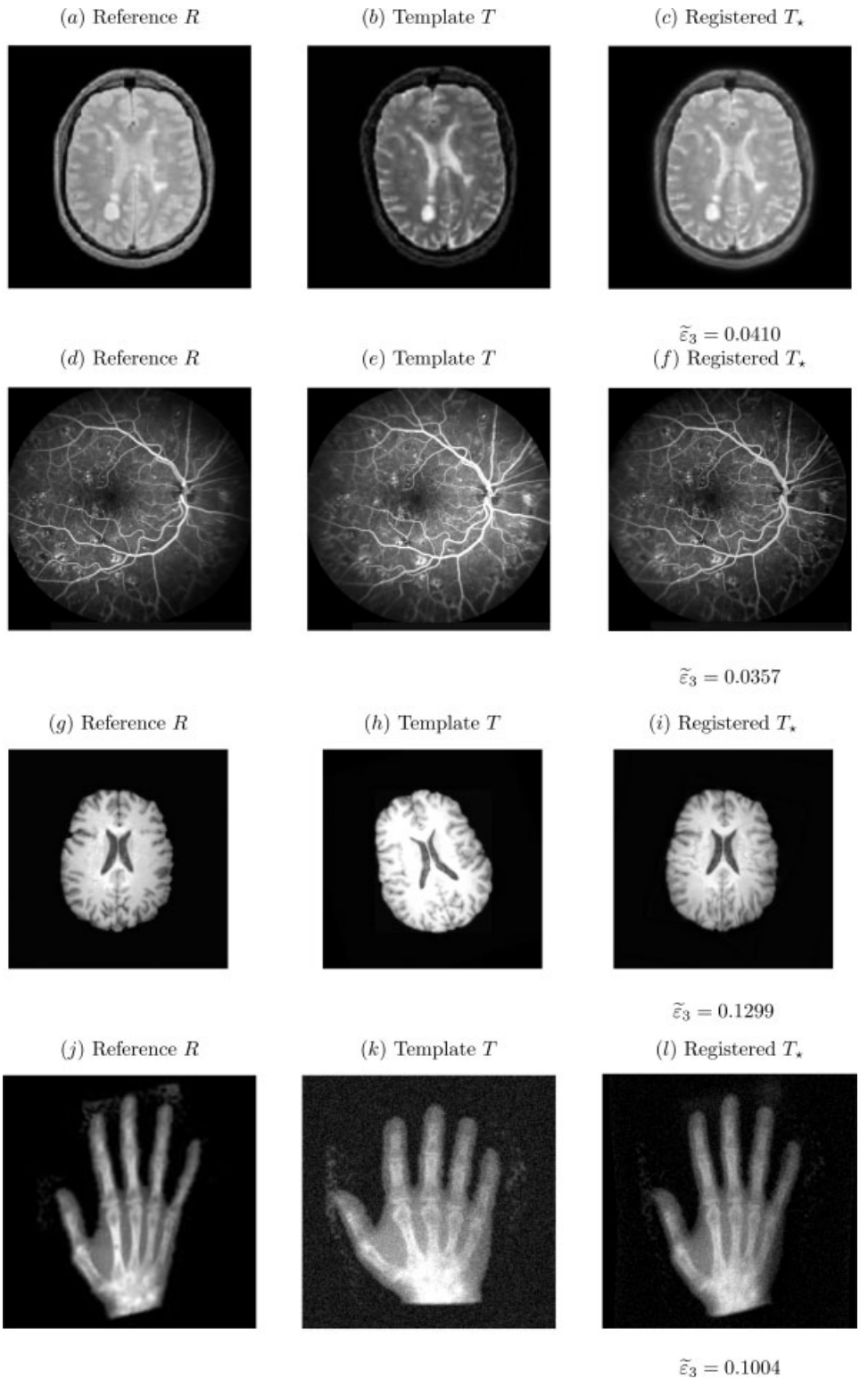

FIG. 8. Four numerical registration results with our new model.

TABLE I. Registration results of Algorithm 2 with the proposed solver in Algorithm 1 for processing four sets of clinical data shown in the first and second columns of Fig. 8.

	Example 1 M/D/WUs	Example 2 M/D/WUs	Example 3 M/D/WUs	Example 4 M/D/WUs
	$\alpha_1 = 10^{-4}, \alpha_1 = 2\alpha_2$	$\alpha_1 = \alpha_2 = 10^{-4}$	$\alpha_1 = 10^{-4}, \alpha_1 = 2\alpha_2$	$\alpha_1 = \alpha_2 = 10^{-4}$
$h = 1/128$	1/0.0322/27	1/0.0264/27	2/0.1183/53	1/0.0914/27
$h = 1/256$	1/0.0381/27	1/0.0301/27	2/0.1242/53	1/0.0953/27
$h = 1/512$	1/0.0410/27	1/0.0357/27	2/0.1299/53	2/0.1004/53
$h = 1/128$	5/0.0321/133	5/0.0263/133	6/0.1182/160	5/0.0913/133
$h = 1/256$	5/0.0381/133	5/0.0300/133	6/0.1241/160	5/0.0952/133
$h = 1/512$	5/0.0409/133	5/0.0356/133	6/0.1298/160	6/0.1003/133

The letters "M," "D," and "WUs" mean the number of MG steps, the relative reduction of dissimilarity ($\widetilde{\varepsilon}_3$), and the work units, respectively. The last three rows are results for dropping the mean of relative residuals to 10^{-4}.

V. NUMERICAL EXPERIMENTS AND RESULTS

To validate and evaluate our new variational registration model (14) and the performance of our Algorithm 2, we first performed a series of tests to verify the model effectiveness and to compare with previous work. Second, we tested with respect to different resolutions. In all registration problems, the bilinear interpolation was used to compute the transformed template image $T(\boldsymbol{u})$ and we stared our MG algorithm with $\nu_1 = \nu_2 = \text{PCGSiter} = 10$, $\beta_1 = 1$, $\beta_2 = 10^{-2}$, $\boldsymbol{u}^{[0]} = 0$, and $c^{[0]} = 0$.

A. Qualify And Comparison of Registration

In the first test, we evaluated the robustness of the proposed registration model (14) or (19) for the cases where the required geometric and intensity transformations are very complex. Shown in Fig. 8 are results from four clinical cases. In each case, the reference R and the template T are from different view points and times. Shown across each row are the reference R and template T and the registered image $T_\star(\boldsymbol{u})$. Even in the presence of significant intensity variations, the registered images $T_\star(\boldsymbol{u})$ by our new model (19) are in good agreement with the reference and show good qualitative registration results.

In the second test, we used the problem in Fig. 8 (d,e) to compare two regularizers $\mathcal{K}(c)$ (this work) and $\mathcal{R}^{L_2}(c)$ (as used in Ref. [13]), with results shown in Fig. 9. Although the registered images by two regularization techniques for c are almost identical by error comparisons, the registration results in Fig. 9 (c,d; after zooming in) confirm that our new model is better for this more challenging problem (where very accurate results are required for clinical image analysis); compare the white arrow locations.

Finally, shown in Fig. 10 are more comparative results from the third test for the problem in Fig. 8(g,h). They indicate that our PDE-based registration model (19) is more robust than that of the previous work (8) of Ref. [13]. Here, $\mathcal{D} = \mathcal{D}_{\text{II}}^{\text{SSD}}$, $(\mu, \lambda, \alpha_1, \alpha_2) = (1.00, 1.00, 0.10, 0.05)$ and the Dirichlet boundary condition $u_l = 0$ $(l = 1, 2)$ were used throughout this test.

B. MG Performance

As is well known, the main property of MG algorithms is that their convergence does not depend on an increasing sequence of resolutions (or a decreasing mesh parameter h). Therefore, in the next round of tests, we designed our experiments to investigate this property.

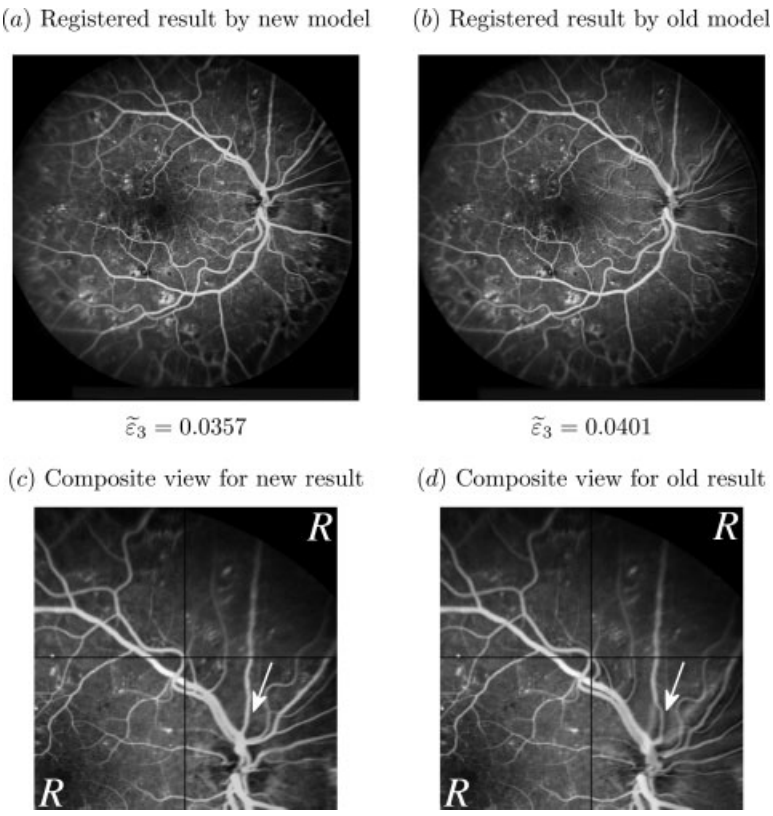

FIG. 9. Numerical results for the second problem shown in Fig. 8(d, e) by two regularization techniques for c. Left to right: results by $\mathcal{K}(c)$ and $\mathcal{R}^{L_2}(c)$. Top to bottom: registered images and composite views in the middle right regions between R and $T_\star(\boldsymbol{u})$. As shown in (c), the intensity variations of $T_\star(\boldsymbol{u})$ in the bottom right region (at the white arrow location) by $\mathcal{K}(c)$ is well matched, compared with those of (d) by $\mathcal{R}^{L_2}(c)$.

To do this, we resolve four registration problems of medical data as shown in the first and second columns in Fig. 8 and started the registration processes with $h = 1/128, 1/256, 1/512$. Here, we define a work unit used in measuring computational complexity as the work of performing a smoother or relaxation step on the finest grid defined as follows:

$$\begin{aligned} 1\ \text{WU} \ &= (\text{cost of discretizating and constructing the linearized system per grid point} \\ &\quad + \ \text{cost of PCGS updating per grid point})N \quad (\text{if } N \text{ is the number of grid points}). \end{aligned}$$

Thus, a work unit in performing one step of our smoother can be estimated by

$$1\ \text{WU} \ = [177 + 284(\text{PCGSiter})]N$$

where each grid point in the linearized system (6×6) given in (48) is solved by the Gaussian elimination method, which has the cost of $\frac{(6)^3}{3} + \frac{(6)^2}{2} - \frac{5(6)}{6}$ additions and $\frac{(6)^3}{3} + (6)^2 - \frac{(6)}{3}$

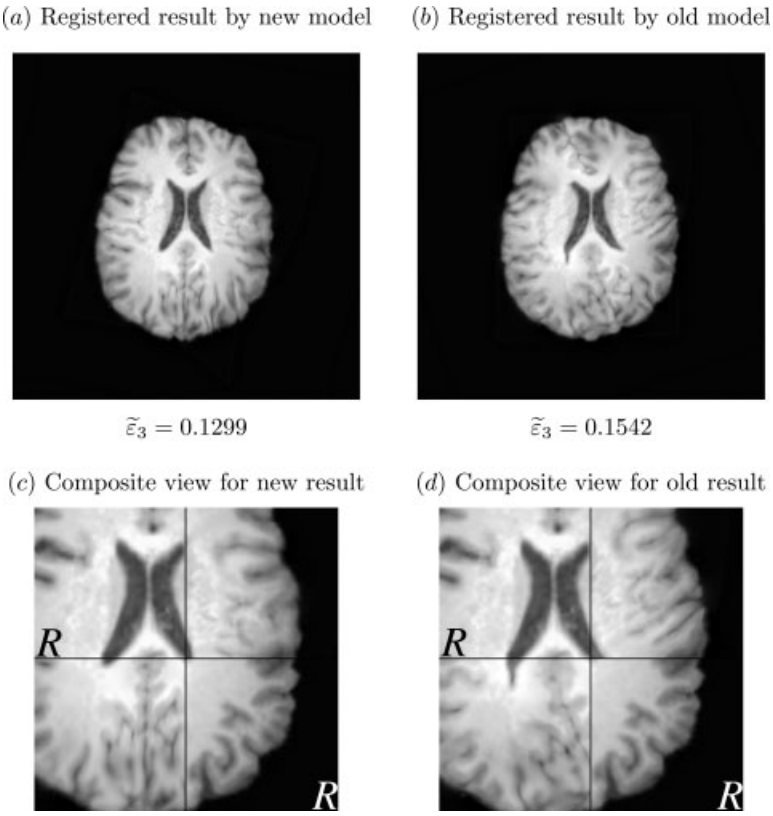

FIG. 10. Numerical results for the third problem shown in Fig. 8(g, h) by two PDE-based registration models. Left to right: results by our model (22) and model (8) of Ref. [13]. Top to bottom: registered images and composite views between R and $T_\star(\boldsymbol{u})$. The top right and bottom left regions of $T_\star(\boldsymbol{u})$ in (c) by our model are well-registered with the adjacent regions of R, compared with those of (d) by Ref. [13].

multiplications. Therefore, the total cost of one V-cycle used L coarse grids can be estimated as follows:

$$\text{V-cycle cost} = (\nu_1 + \nu_2)[177 + 284(\text{PCGSiter})]N \sum_{k=0}^{L} (1/4)^k < \frac{4}{3}(\nu_1 + \nu_2)\ \text{WUs}.$$

Here, we have ignored the cost of interpolation and restriction procedures as well as the cost of residual correction procedure because they are relatively small compared with smoothing procedures. Recall that ν_1, ν_2, and PCGSiter denote respectively the number of presmoothing and postsmoothing and PCGS steps.

The numerical results are reported in Table I where one can see three quantities: the numbers of MG cycles "M"; the relative reduction of dissimilarity $\mathrm{D} = \widetilde{\varepsilon}_3$; and the work units "WUs".

As expected from a MG technique, Table I shows that our MG algorithm is h-independent, that is, scalable. Moreover, it took only one or two MG steps to solve the registration problems and reduce the dissimilarities between the reference and registered images by more than 85% for all problems we tested.

The authors thank the reviewers and Dr Colin Baker (Clatterbridge Centre for Oncology, UK) for making helpful suggestions to improve the paper.

VI. CONCLUSIONS

In a multimodality problem where two images to be registered have different intensity ranges, the standard model of the SSD is not sufficient as it can fail to register such images. In this article, we have proposed an improved model of the SSD using homogenization-based correction ideas and incorporating curvature regularizers for both the displacement field and intensity correction. Although the results of our model show a major improvement over related models, the underlying Euler–Lagrange system consisting of more complex nonlinear PDEs is less amenable to efficient solution than models without intensity corrections. To this end, a MG algorithm using a new PDFP smoother is developed to speed up the solution. Numerical tests confirm that both of our registration model and MG algorithm are reliable in reducing registration errors and providing visually pleasing results for practical applications.

References

1. A. Madabhushi and J. K. Udupa, New methods of MR image intensity standardization via generalized scale, Med Phys 33 (2006), 3426–3434.
2. A. Madabhushi, J. K. Udupa, and G. Moonis, Comparing MR image intensity standardization against tissue characterizability of magnetization transfer ratio imaging, J Magn Reson Imaging 24 (2006), 667–675.
3. J. V. Manjón, J. J. Lull, J. Carbonell-Caballero, and G. García-Martí, A nonparametric MRI inhomogeneity correction method, Med Image Anal 11 (2007), 336–345.
4. L. G. Nyúl and J. K. Udupa, On standardizing the MR image intensity scale, Magn Reson Med 42 (1999), 1072–1081.
5. Y. Zhuge, J. K. Udupa, J. Liu, and P. K. Saha, Image background inhomogeneity correction in MRI via intensity standardization, Comput Med Imaging Graph 33 (2009), 7–16.
6. E. D'Afostino, F. Maes, D. Vandermeulen, and P. Suetens, A viscous fluid model for multimodal non-rigid image registration using mutual information, Med Image Anal 7 (2003), 565–575.
7. E. Haber and J. Modersitzki, Intensity gradient based registration and fusion of multi-modal images, Med Image Comput Comput Assist Interv 3216 (2006), 591–598.
8. X. Lu, S. Zhang, H. Su, and Y. Chen, Mutual information-based multimodal image registration using a novel joint histogram estimation, Comput Med Imaging Graph 32 (2008), 202–209.
9. J. Modersitzki, FAIR: flexible algorithms for image registration, SIAM, Philadelphia, PA, 2009.
10. H. Nam, R. A. Renaut, K. Chen, H. Guo, and G. E. Farin, Improved inter-modality image registration using normalized mutual information with coarse-binned histograms, Commun Numer Methods Eng 25 (2009), 583–595.
11. J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, Image registration by maximization of combined mutual information and gradient information, IEEE Trans Med Imaging 19 (2000), 809–814.
12. D. Rueckert, M. J. Bowden, D. J. G. Hill, and D. J. Hawkes, Image template matching using mutual information and NP-windows, Med Imaging 2 (2000), 438–447.
13. J. Modersitzki and S. Wirtz, Combining homogenization and registration, Lect Notes Comput Sci 4057 (2006), 257–263.

14. A. Guimond, A. Roche, N. Ayache, and J. Meunier, Three-dimensional multimodal brain wrapping using the demons algorithm and adaptive intensity corrections, IEEE Trans Med Imaging 20 (2001), 58–69.

15. S. Periaswamy and H. Farid, Elastic registration in the presence of intensity variations, IEEE Trans Med Imaging 22 (2003), 865–874.

16. W. Ou and C. Chefd'Hotel, Polynomial intensity correction for multimodal image registration, Proceedings of the Sixth IEEE international conference on Symposium on Biomedical Imaging: From Nano to Macro, Boston, MA, USA, 2009, pp. 939–942.

17. E. R. Arce-Santana and A. Alba, Image registration using Markov random coefficient and geometric transformation fields, Pattern Recognit 42 (2009), 1660–1671.

18. J. P. Thirion, Image matching as diffusion process: an analogy with Maxwell's demons, Med Image Anal 2 (1998), 243–260.

19. J. L. Marrroquin, B. C. Vemuri, S. Botello, F. Calderon, and A. Fernandez-Bouzas, An accurate and efficient Bayesian model for automatic segmentation of brain MRI, IEEE Trans Med Imaging 21 (2002), 934–945.

20. N. Badshah and K. Chen, Multigrid method for the Chan–Vese model in variational segmentation, Commun Comput Phys 4 (2008), 294–316.

21. N. Badshah and K. Chen, On two multigrid algorithms for modelling variational multiphase image segmentation, IEEE Trans Image Process 18 (2009), 1097–1106.

22. C. Brito-Loeza and K. Chen, Multigrid method for a modified curvature driven diffusion model for image inpainting, J Comput Math 26 (2008), 856–875.

23. A. Bruhn and J. Weickert, Multigrid platform for real-time motion computation with discontinuity-preserving variational methods, Int J Conf Violence 70 (2006), 257–277.

24. T. F. Chan and K. Chen, On a nonlinear multigrid algorithm with primal relaxation for the image total variation minimization, J Numer Algorithms 41 (2006), 387–411.

25. N. Chumchob and K. Chen, A variational approach for discontinuity-preserving image registration, East-West J Math, Special Volume:266–282, 2010.

26. K. Chen, Y. Dong, and M. Hintermüller, A nonlinear multigrid solver with line gauss-seidel-semismooth-newton-smoother for the fenchel-pre-dual in total variation based image restoration, IFB-Report, Vol. 37, IMSC (University of Graz, Austria), Inverse Problem and Imaging 5 (2011), 323–339.

27. C. Frohn-Schauf, S. Henn, and K. Witsch, Multigrid based total variation image registration, Comput Visual Sci 11 (2008), 101–113.

28. S. Gao, L. Zhang, H. Wang, R. de Crevoisier, D. D. Kuban, R. Mohan, and L. Dong, A deformable image registration method to handle distended rectums in prostate cancer radiotherapy, Med Phys 33 (2006), 3304–3312.

29. E. Haber, R. Horesh, and J. Modersitzki, Numerical optimization for constrained image registration, Numer Linear Algebra Appl 17 (2010), 343–359.

30. E. Haber and J. Modersitzki, A multilevel method for image registration, SIAM J Sci Comput 27 (2006), 1594–1607.

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

32. H. Köstler, K. Ruhnau, and R. Wienands, Multigrid solution of the optical flow system using a combined diffusion- and curvature-based regularizer, Numer Linear Algebra Appl 15 (2008), 201–218.

33. D. Zikic, W. Wein, and A. Khamene, Fast deformable registration of 3D-ultrasound data using a variational approach, Lect Notes Comput Sci 4190 (2006), 915–923.

34. J. Modersitzki, Numerical methods for image registration, Oxford University Press, Oxford, 2004.

35. B. Fischer and J. Modersitzki, A unified approach to fast image registration and a new curvature based registration technique, Linear Algebra Appl 380 (2004), 107–124.

36. S. Henn, A full curvature based algorithm for image registration, J Math Imaging Vis 24 (2006), 195–208.

37. C. Brito-Loeza and K. Chen, Multigrid algorithm for high order denoising, SIAM J Imaging Sci 3 (2010), 363–389.

38. 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), 1964–1997.

39. N. Chumchob and K. Chen, A robust affine image registration method, Int J Numer Anal Mod 6 (2009), 311–334.

40. T. F. Chan and P. Mulet, On the convergence of the lagged diffusivity fixed point method in total variation image restoration, SIAM J Numer Anal 36 (2007), 354–367.

41. T. F. Chan and J. H. Shen, Image Processing and analysis: variational, PDE, wavelet, and stochastic methods, SIAM, Philadelphia, PA, 2005.

42. J. Savage and K. Chen, An improved and accelerated nonlinear multigrid method for total-variation denoising, Int J Comput Math 82 (2005), 1001–1015.

43. C. R. Vogel and M. E. Oman, Iterative methods for total variation denoising, SIAM J Sci Comput 17 (1996), 227–238.

44. C. R. Vogel, Computational methods for inverse problems, SIAM, Philadelphia, PA, 2002.

45. W. L. Briggs, V. E. Henson, and S. F. McCormick, A multigrid tutorial, 2nd Ed., SIAM, Philadelphia, PA, 2000.

46. W. Hackbusch, Multi-grid methods and applications, Springer-Verlag, Berlin, 1985.

47. U. Trottenberg, C. Oosterlee, and A. Schuller, Multigrid, Elsevier, Academic Press, Bodmin, Cornwall, 2001.

48. P. Wesseling, Multigrid methods, R. T. Edwards, Philadelphia, PA, 2004.

49. R. Wienands and W. Joppich, Practical fourier analysis for multigrid method, Chapman & Hall/CRC Press, Boca Raton, FL, 2005.

50. T. F. Chan and K. Chen, An optimization-based multilevel algorithm for total variation image denoising, Multiscale Model Simul 5 (2006), 615–645.

51. S. Hamilton, M. Benzi, and E. Haber, New multigrid smoothers for the Oseen problem, Numer Linear Algebra Appl 17 (2010), 556–576.

52. B. Seynaeve, E. Rosseel, B. Nicolaï, and S. Vandewalle, Fourier mode analysis of multigrid methods for partial differential equations with random coefficients, J Comput Phys 224 (2007), 132–149.

53. K. Chen, Matrix preconditioning techniques and applications, Cambridge University Press, UK, 2005.