Variational image registration by a total fractional-order variation model ☆

Jianping Zhang [a,b], Ke Chen [a,*]

[a] *Centre for Mathematical Imaging Techniques and Department of Mathematical Sciences, The University of Liverpool, Peach Street, Liverpool L69 7ZL, United Kingdom*
[b] *School of Mathematics and Computational Science, Xiangtan University, Xiangtan, Hunan 411105, PR China*

ARTICLE INFO

ABSTRACT

In this paper, a new framework of nonlocal deformation in non-rigid image registration is presented. It is well known that many non-rigid image registration techniques may lead to unsteady deformation (e.g. not one to one) if the dissimilarity between the reference and template images is too large. We present a novel variational framework of the total fractional-order variation to derive the underlying fractional Euler–Lagrange equations and a numerical implementation combining the semi-implicit update and conjugate gradients (CG) solution to solve the nonlinear systems. Numerical experiments show that the new registration not only produces accurate and smooth solutions but also allows for a large rigid alignment, the evaluations of the new model demonstrate substantial improvements in accuracy and robustness over the conventional image registration approaches.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

One of the most important tasks in computer vision and image processing is registration, aiming to find a geometrical transformation that aligns points in one view of an object with corresponding points in another view of that object or another object, i.e., realigns two images – the reference and template images. Nowadays, image registration has played an important role in different applications, such as remote sensing, medicine and computer vision. Especially in medical diagnosis [1–4], for example, the efficient implementation of the automating medical diagnosis with the aid of computers should depend on reliable registration methods.

In addition to simple parameter based methods [5], the optical flow based approach is an early variational method, aiming to recover the displacement field between two frames of a video sequence which are taken at different times at every voxel position, so local Taylor series approximations of the image signal and the partial derivatives with respect to the spatial and temporal coordinates are used to calculate the motion between two images.

During the last decades, to realize image registration, a great number of variational approaches in the purpose of minimizing the similarity measures have been proposed. The similarity measures are used to quantify the degree of similarity

* Corresponding author.
E-mail addresses: jpzhang@xtu.edu.cn (J. Zhang), k.chen@liverpool.ac.uk (K. Chen).
URL: http://www.liv.ac.uk/~cmchenk (K. Chen).

between intensity patterns within two images. Since the underling problem is in general ill-posed in the sense of Hadamard, therefore, how to effectively minimize the similarity measures becomes a fundamental task in image sciences.

Regularizing ensures that the resulting well-posed problem admits a solution. In Tikhonov framework, the cost energy functional minimized in the registration model is a combination of the image similarity and the regularizing penalty functional. On one hand the choice of an image similarity measure depends on the modality of the images to be registered [6], including single-modality and multi-modality methods. Single-modality methods tend to register images in the same modality acquired by the same scanner type [7,8], while multi-modality registration methods tend to register images acquired by different scanner types. Common examples of image similarity measures include cross-correlation, mutual information and sum of squared intensity differences (SSD) [7,8]. Mutual information and normalized mutual information are the two of most popular image similarity measures for registration of multimodality images, while cross-correlation and SSD are commonly used for registration of images in the same modality. Image registration algorithms can be also classified into intensity-based and feature-based [6]. Intensity-based methods compare intensity patterns in images via correlation metrics, while feature-based methods find correspondence between image features such as points, lines, and contours [6].

On the other hand, image registration algorithms can also be classified according to the transformation models being used to relate the template (target) image space to the reference image space. The first category of transformation models refers to linear/affine transformations, which include rotation, scaling, translation [7]. Linear transformations are global in nature, thus, they cannot model local geometric differences between images [6]. The second category of transformations allows 'elastic' or 'nonrigid' transformations. These transformations are capable of locally warping the template image to align with the reference image. Nonrigid transformations include radial basis functions [6], physical continuum models [8–12] and large deformation models (diffeomorphisms) [13]. In all cases, it is preferable to choose transformations that have physical meaning, but in some cases, the choice is made on the basis of convenient mathematical properties. However, large local and global deformations may occur and must be taken into account.

Over the last decade, it has been demonstrated that many systems in science and engineering can be modeled more accurately by employing fractional-order rather than integer-order derivatives [14–16], and many methods are developed to solve the fractional systems [17–22]. Not all of these results have been considered for imaging applications. Recently, there have been several works involving discrete forms of an α-order derivative proposed to tackle the image restoration problem [23–30] and the image inpainting problem [31]. However much fewer works employing partial fractional α-order derivatives are applied to the image registration problem. Melbourne et al. [32] used fractional differentiation (differentiation to non-integer order) to design new gradients of image intensities for enhancing image registration performance to directly register image gradients. Garvey et al. in [33] proposed a nonrigid registration algorithm that involves directly and rapidly solving a discretized fractional PDE modeling super-diffusive processes in nonrigid image registration. The proposed algorithm yields lower average deformation errors than standard diffusion-based registration through registration experiments on breast MR imagery with simulated biomechanical deformations. In [34], a regularization term based on fractional order derivatives is introduced but the problem is solved in the frequency domain of the minimizing energy functional via the Euler–Lagrange equations. In [35], medical image registration was studied in the domain of fractional Fourier transform. These earlier works have suggested and illustrated that fractional order derivatives may be effective regularizers for image registration applications.

The contributions of this work are the following

i). We propose a new nonlocal deformation model with the total fractional-order variation regularizer in non-rigid image registration in a continuous setting. Due to the nonlocal field theories of fractional derivative, the new registration can produce accurate and steady smooth deformation.

ii). We establish better and more rigorous theories for applications of the total fractional-order variation to image inverse problems. To apply the total fractional-order variation regularization with fractional order derivative to variational image inverse problems, we analyze properties of the total fractional-order variation and its fractional integration by parts formulas from variational principles. We derive the Euler–Lagrange equation in suitable function spaces.

iii). We present a new numerical scheme combining the semi-implicit update, discretization matrix approximation and CG iterative solution.

Our work will facilitate future applications of α-order variation based regularizer to other imaging problems where regularization is required.

The rest of the paper is organized as follows. Section 2 first reviews the basic image registration problem, the demons algorithm, several variational models about image registration, definitions and basic properties of the fractional order derivative which help us to understand the differences between integer and fractional derivatives. In Section 3 we first discuss definition and properties of the total fractional-order variation which generates TV regularization with integer derivative. Then a total fractional-order variation image registration model with nonlocal property is considered, we should discuss the derivation of Euler–Lagrange equation and boundary condition. Before ending this section the study of discretization of Euler–Lagrange equation and efficient numerical schemes are developed. Experimental results are shown in Section 4, and the paper is concluded with a summary in Section 5.
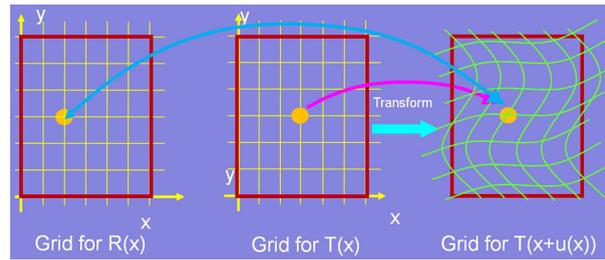
Fig. 1. Non-rigid registration: $R(x)$ – reference image; $T(x)$ – template image. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

2. The image registration modeling and fractional derivatives

Below we shall first introduce the basic problem of registration, notation and the variational framework.

2.1. Preliminaries

Let Ω be a bounded domain in $\mathbb{R}^d$, without loss of generality we assume $\Omega = [0, a]^d$, where $d \in \mathbb{N}$ denotes the spatial dimension of the images. The known *Reference* and *Template* images of the same object taken at different times can be defined respectively by Lipschitz functions $R(x) \in I$ and $T(x) \in I$, where $I = [0, 255]$ is the range of gray image intensities, especially we also assume that $T(x + \boldsymbol{u})$ is a compactly supported function w.r.t. the displacement field $\boldsymbol{u}$. Each view in Fig. 1 that is involved in a registration will be referred to a coordinate system, which defines a space grid for that view. Our definition of registration is based on geometrical transformation, which is mappings of points from the grid of one view to the grid of a second view. The transformation applied to a point (*see* the yellow point in the left picture of Fig. 1) in grid for reference image R represented by the column vector $x \in \Omega$ produces a transformed point $\phi(\boldsymbol{u}(x)) \in \mathbb{R}^d$,

$$\phi(\boldsymbol{u}(x)) = x + \boldsymbol{u}(x).$$

If the point $\phi(\boldsymbol{u}(x))$ corresponds to x, then a successful registration will make $T(\phi(\boldsymbol{u}(x)))$ equal, or approximately equal, to $R(x)$ (*see* two yellow points in the left and right pictures of Fig. 1). The above ϕ is called as the additive structure of the displacement field; a Lie group structure on diffeomorphisms is also used to model the spatial transformation of coordinates [13]. The purpose of registration is to seek a displacement field $\boldsymbol{u} : \Omega \to \Omega$ of image coordination such that the transformed template $T(x + \boldsymbol{u}(x)) = R(x)$ approximately. Once the corresponding location $\phi(\boldsymbol{u}(x)) = x + \boldsymbol{u}(x)$ is calculated for each spatial location $x \in \Omega$, an image interpolation is required to assign the image intensity values for the transformed template $T(\boldsymbol{u})$ at non-grid locations within image boundaries. Now the question is how to find such a mapping $\boldsymbol{u} = (u_1, \ldots, u_d)^t$. As we have said, a typical approach is the minimization of a suitable distance measure $D(\boldsymbol{u})$. For example, assume the image intensities of $R(x)$ and $T(x)$ are comparable (i.e., in a monomodal registration scenario), the so-called sum of squared differences measure is

$$D(\boldsymbol{u}) := \frac{1}{2} \int_{\Omega} (T(x + \boldsymbol{u}(x)) - R(x))^2 dx. \tag{1}$$

The task of finding a reasonable transformation such that a transformed version of a template image is similar to a reference image is to solve the minimization problem of the similarity measure: $\min_{\boldsymbol{u}} D(\boldsymbol{u})$.

2.2. Demon's algorithm to image registration

The demons algorithm proposed firstly by Thirion [36] provides a very efficient registration scheme with simple implementation and linear computational complexity for nonrigid image registration problem, which is considered as a popular technique for fast intensity-based registration, particularly within a given modality where image intensity values (i.e., voxel values) are consistent, and has been applied to a range of applications [13,37].

The standard demons algorithm can be considered as an approximation of a second-order gradient descent of the sum of square of intensity differences metric $D(\boldsymbol{u})$. Minimization of $D(\boldsymbol{u})$ is very close to a simple mean squared error image registration problem whose goal is to find an optimal field $\boldsymbol{u}$. In dealing with a least-square problem, the methods they used in their work rely on a linearization of $T(x + \boldsymbol{u}(x)) - R(x)$ in $D(\boldsymbol{u})$ and are based on Gauss–Newton-like approaches. The demons algorithm approximately solves the diffusion registration problem by successively estimating force vectors that drive the deformation toward alignment and smoothing the force vectors by Gaussian convolution.

Let $p^x(\boldsymbol{u}) = T(x + \boldsymbol{u}(x)) - R(x)$ be the intensity difference at point x and assume that the following linearization

$$p^x(\boldsymbol{u}) = p^x(\boldsymbol{u}^k) + \nabla_{\boldsymbol{u}} p^x(\boldsymbol{u}^k) \cdot (\boldsymbol{u} - \boldsymbol{u}^k) + O(\|\boldsymbol{u} - \boldsymbol{u}^k\|^2)$$

$$= T(x + \boldsymbol{u}^k) - R(x) + \nabla T(x + \boldsymbol{u}^k) \cdot (\boldsymbol{u} - \boldsymbol{u}^k) + O(\|\boldsymbol{u} - \boldsymbol{u}^k\|^2)$$

is available. The above can be used to rewrite the correspondence energy in (1) as follows:

$$D(\boldsymbol{u}) \approx \tilde{D}(\boldsymbol{u}) := \frac{1}{2} \int_{\Omega} (T(x + \boldsymbol{u}^k) - R(x) + \nabla T(x + \boldsymbol{u}^k) \cdot (\boldsymbol{u} - \boldsymbol{u}^k))^2 dx, \tag{2}$$

with its minimizer $u^*(x)$ given by linear systems

$$A \cdot (\boldsymbol{u}(x) - \boldsymbol{u}^k) = -(T(x + \boldsymbol{u}^k) - R(x))(\nabla T(x + \boldsymbol{u}^k))^t$$

where $A := (\nabla T(x + \boldsymbol{u}^k))^t \nabla T(x + \boldsymbol{u}^k)$ is a 2×2 symmetric positive semi-definite matrix with $rank(A) \leq 1$. Direct optimization of (2) will lead to an ill-posed problem with unstable and non-smooth solutions. To avoid this and possibly add a priori knowledge, for example, a least square solution is obtained from a well-posed problem after adding a simple regularizer to (2)

$$\min_{\boldsymbol{u}} \left\{ \tilde{J}(\boldsymbol{u}) = \frac{1}{2} \int_{\Omega} (T(x + \boldsymbol{u}^k) - R(x) + \nabla T(x + \boldsymbol{u}^k) \cdot (\boldsymbol{u} - \boldsymbol{u}^k))^2 + \frac{\sigma}{2} \|\boldsymbol{u} - \boldsymbol{u}^k\|^2 dx \right\}.$$

Finally at every point x we get transformation $\boldsymbol{u}(x)$ such that

$$(A + \sigma I)(\boldsymbol{u}(x) - \boldsymbol{u}^k) + (T(x + \boldsymbol{u}^k) - R(x))(\nabla T(x + \boldsymbol{u}^k))^t = 0,$$

which is solved by using a Sherman–Morrison formula $(\omega \omega^t + \sigma I)^{-1} = \frac{1}{\sigma}[I - \frac{\omega \omega^t}{\sigma + \|\omega\|^2}]$

$$\boldsymbol{u}(x) = \boldsymbol{u}^k(x) - \frac{(T(x + \boldsymbol{u}^k) - R(x))(\nabla T(x + \boldsymbol{u}^k))^t}{\sigma + \|\nabla T(x + \boldsymbol{u}^k)\|^2}.$$

Comparing with global transformation cases (e.g., rigid body transformations and variational registration), the approximations at every node x given by demon's algorithm are independent from each other. This greatly simplifies the minimization of $D(\boldsymbol{u})$ by splitting it into very simple systems for each node. In [38], a variant of the demons deformation field which uses symmetric gradient information from both the *reference* and *template* images is proposed in order to perform better than the conventional displacement systems.

2.3. Variational framework

Now return to the problem (1) which is generally ill-posed in the sense of Hadamard. It becomes necessary to impose a constraint on the solution $\boldsymbol{u}$ via a deformation regularizer $S(\boldsymbol{u})$ for penalizing unwanted and steady solutions, where we incorporate a priori knowledge such as smoothness of the desired solution, the error level in two images or the statistical properties of the transforming process. As a consequence, the image registration problem can be posed as a minimization problem of the joint energy functional given by

$$\min_{\boldsymbol{u}} \left\{ E(\boldsymbol{u}) := S(\boldsymbol{u}) + \lambda D(\boldsymbol{u}) = S(\boldsymbol{u}) + \frac{\lambda}{2} \int_{\Omega} (T(x + \boldsymbol{u}(x)) - R(x))^2 dx \right\} \tag{3}$$

where $\lambda > 0$ is the regularization parameter that compromises similarity and regularity, and the term $S(\boldsymbol{u})$ could also be of the geometric properties of the displacement field $\boldsymbol{u}(x)$ which could be the first or second derivative or mean curvature, etc. Observe that given λ, problem (3) becomes a nonlinear optimization problem that we can solve with the first variation and nonlinear techniques for medium and large-scale problems. However, determining an optimal value for the regularization parameter λ can be as difficult as the original problem and most of the methods currently available require the solution of several problems of type (3) for different values of λ; these approaches might be very expensive in large-scale setting. Throughout the paper we assume that the best λ for every model is manually given.

Obviously, the choice of the deformation regularizer $S(\boldsymbol{u})$ is very crucial for effective registration. Different choices of $S(\boldsymbol{u})$ lead not only to different deformation fields but also to different coupled nonlinear systems. Here, the regularizer $S(\boldsymbol{u})$ is reviewed briefly.

Linear elastic regularizer. In [9,12], a linear elastic regularizer is used for the realignment of mono-modal image registration, the corresponding regularization energy is given by functional

$$S(\boldsymbol{u}) := \int_{\Omega} \frac{\mu}{2} (\text{Tr}(E))^2 + \nu \, \text{Tr}(E^2) dx,$$

where $\text{Tr} = trace$ and $E = \frac{1}{2}(\nabla \boldsymbol{u}^T + \nabla \boldsymbol{u})$ is a linear approximation of a strain tensor, μ and ν are the Lame constants (known for a given elastic material). This model measuring the energy of the elastic deformation is isotropic in the directions and is neutral with respect to translations and rotations but penalizes these transformations by the Dirichlet boundary conditions [39]. Therefore, this scheme is very attractive for high-resolution applications.

Diffusion regularizer. Diffusion is the flow of molecules (energy) from a place of high concentration to another of low concentration. Diffusion processes strive to equilibrate the concentration differences in the system whilst preserving the total mass of the system. The diffusion-based image registration by Fischer and Modersitzki [10] chose the following diffusion regularizer

$$S(\boldsymbol{u}) := \frac{1}{2} \int_\Omega \left(\sum_{\ell=1}^d |\nabla u_\ell|^2 \right) dx.$$

Linear curvature regularizer. The linear curvature registration [11] is based on the following regularizer

$$S(u) := \frac{1}{2} \int_\Omega \left(\sum_{\ell=1}^d (\Delta u_\ell)^2 \right) dx.$$

The reason for this particular choice is twofold. Firstly the integral might be viewed as an approximation to the curvature of the ℓ-th component of the displacement field and therefore does penalize oscillations. Secondly, $S(\boldsymbol{u})$ has a non-trivial kernel containing harmonic functions and in particular affine linear transformations.

Mean curvature regularizer. The mean curvature registration technique [8] is based on the following regularizer

$$S(\boldsymbol{u}) := \frac{1}{2} \int_\Omega \left(\sum_{\ell=1}^d (\nabla \cdot \frac{\nabla u_\ell}{\sqrt{1 + |\nabla u_\ell|^2}})^2 \right) dx.$$

This model does not require an additional affine linear pre-registration step for being successful, in contrast to many other non-linear registration techniques, including the elastic matching, the fluid matching, and the diffusion matching. However, folding of deformations (20) can be a problem.

2.4. Fractional-order derivatives

Fractional calculus is the branch of mathematics that generalizes the integer order derivatives and integrals of a function to non-integer order. Fractional derivatives have been used to model physical and engineering processes, which could be more efficiently described by fractional differential equations, as seen in this special issue. The fractional order derivative has a long history, which unifies differential and integral operators into one differential-integral operator. As we know, the integer order derivatives are both unique and local, while the fractional derivatives and their inverses are generally non-local.

Several definitions have been proposed to describe a fractional order derivative [16,40]; we shall present three of them below. For a systematic presentation of the mathematics, a fractional α order derivative is denoted as function operator $D_{[a,x]}^\alpha$, where a and x are the bounds of the integral, $[a, x]$ defines a 1D computational domain, and $0 \le \ell := n - 1 < \alpha < n$. For $\alpha > 1$, we consider $\alpha = n - (n - \alpha)$ with $0 < (n - \alpha) < 1$.

1. **Riemann–Liouville (RL) definitions.** The left and right-sided RL derivatives of order α of a function $f(x)$ are given as follows:

$$D_{[a,x]}^\alpha f(x) = \frac{1}{\Gamma(n-\alpha)} \left(\frac{d}{dx} \right)^n \int_a^x \frac{f(\tau)d\tau}{(x-\tau)^{\alpha-n+1}} \quad \text{and} \quad D_{[x,b]}^\alpha f(x) = \frac{1}{\Gamma(n-\alpha)} \left(-\frac{d}{dx} \right)^n \int_x^b \frac{f(\tau)d\tau}{(\tau-x)^{\alpha-n+1}}.$$

2. **Grünwald–Letnikov (GL) definitions.** The left and right-sided GL derivatives are defined by

$$^G D_{[a,x]}^\alpha f(x) = \lim_{h \to 0} \frac{\sum_{j=0}^{[\frac{x-a}{h}]} (-1)^j C_\alpha^j f(x-jh)}{h^\alpha} \quad \text{and} \quad ^G D_{[x,b]}^\alpha f(x) = \lim_{h \to 0} \frac{\sum_{j=0}^{[\frac{b-x}{h}]} (-1)^j C_\alpha^j f(x+jh)}{h^\alpha}.$$

3. **Caputo (C) definitions.** The left and right-sided Caputo derivatives are defined by

$$^C D_{[a,x]}^\alpha f(x) = \frac{1}{\Gamma(n-\alpha)} \int_a^x \frac{f^{(n)}(\tau)d\tau}{(x-\tau)^{\alpha-n+1}} \quad \text{and} \quad ^C D_{[x,b]}^\alpha f(x) = \frac{(-1)^n}{\Gamma(n-\alpha)} \int_a^x \frac{f^{(n)}(\tau)d\tau}{(\tau-x)^{\alpha-n+1}}.$$

The spatial Riesz-type fractional α-order derivative which is considered as a Riesz-type potential is defined as a half-sum of the left and right-sided derivatives of three definitions for function $f(x)$:

$$^\Upsilon D_{[a,b]}^\alpha f(x) = \frac{1}{2} \left(^\Upsilon D_{[a,x]}^\alpha f(x) + (-1)^n {}^\Upsilon D_{[x,b]}^\alpha f(x) \right),$$

where Υ takes R, RG and RC for Riesz–RL, Riesz–GL and Riesz–Caputo fractional derivative respectively. Note that, for $\bar{\alpha} = (n - \alpha)$, the formula

$$D_{[a,x]}^{-\bar{\alpha}} f(x) := \frac{1}{\Gamma(\bar{\alpha})} \int_a^x \frac{f(\tau)d\tau}{(x-\tau)^{1-\bar{\alpha}}} \tag{4}$$

defines the Abel integral or the fractional $\bar{\alpha}$-order integral, used to define the general fractional α-order derivative.

A crucial step in variational problems is to find minima (or maxima) of an energy functional, the usual integration by parts or a respective divergence theorem can be directly used to offer the weak solution of problems with integer order derivatives. In this work we are concerned with the problems of finding a necessary optimality conditions for functionals of calculus of variations based on fractional order derivative. By inserting fractional derivatives into the variational integrals, the interesting integration by parts formulas in [41] is useful to derive the PDE systems for the above total fractional-order variation based image problems. The procedure to derive formulations and the resulting equations for problems defined in terms of those derivatives turn out to be very similar to before.

Lemma 2.1 (*Integration by parts formulas*). (*See [41].*) *For $0 \le n - 1 < \alpha < n$, we have*

$$\int_a^b \xi(x) \cdot D_{[a,x]}^\alpha \eta(x) dx = \int_a^b \eta(x) \cdot {}^C D_{[x,b]}^\alpha \xi(x) dx - \sum_{j=0}^{n-1} (-1)^{n+j} D_{[a,x]}^{\alpha-n+j} \eta(x) \frac{\partial^{n-j-1}\xi(x)}{\partial x^{n-j-1}} \Big|_{x=a}^{x=b}$$

and

$$\int_a^b \xi(x) \cdot D_{[x,b]}^\alpha \eta(x) dx = \int_a^b \eta(x) \cdot {}^C D_{[a,x]}^\alpha \xi(x) dx - \sum_{j=0}^{n-1} D_{[x,b]}^{\alpha-n+j} \eta(x) \frac{\partial^{n-j-1}\xi(x)}{\partial x^{n-j-1}} \Big|_{x=a}^{x=b}.$$

Thus, in the case of the Riesz R–L fractional derivative, one has

$$\int_a^b \xi(x) \cdot {}^R D_{[a,b]}^\alpha \eta(x) dx = (-1)^n \int_a^b \eta(x) \cdot {}^{RC} D_{[a,b]}^\alpha \xi(x) dx + \sum_{j=0}^{n-1} (-1)^{n+j\,R} D_{[a,b]}^{\alpha-n+j} \eta(x) \frac{\partial^{n-j-1}\xi(x)}{\partial x^{n-j-1}} \Big|_{x=a}^{x=b}.$$

Remark 2.1 (*Equivalence*). In the above lemma, one observes that D^α, ${}^C D^\alpha$ are strongly connected. Indeed they are closely related [16,42]: for every $0 < \alpha < n$, if the function $f(x)$ is $(n-1)$-order continuously differentiable and $f^{(n)}(x)$ is integrable in $[a, b]$, the R–L fractional derivative $D_{[a,x]}^\alpha f(x)$ exists and is equivalent to the G–L derivative ${}^G D_{[a,x]}^\alpha f(x)$. Furthermore if $f(x)$ satisfies the homogeneous boundary conditions, the RL and the Caputo derivatives are also equivalent. However, separately, all these fractional derivatives definitions have their advantages and disadvantages. Refer to [42–44] for more details.

A simple transformation $\tilde{f}(x) = f(x) - \left(f(a) + \frac{f(b)-f(a)}{b-a}(x-a) \right)$ to $f(x)$ can also modify the Riesz RL fractional derivative $D_{[a,b]}^\alpha f(x)$ to avoid singularity at $x = a$. As said in [16], the fractional derivative $D_{[0,x]}^\alpha$ for three above definitions has linearity property, i.e., for any $f(x), g(x)$ and $\beta \in (0, 1)$, then one has

$$\Upsilon D_{[a,b]}^\alpha (\beta f(x) + (1-\beta) g(x)) = \beta^\Upsilon D_{[a,b]}^\alpha f(x) + (1-\beta)^\Upsilon D_{[a,b]}^\alpha g(x),$$

where Υ can be one of R, RG and RC for Riesz–RL, Riesz–GL and Riesz–Caputo fractional derivatives. Such a property is very useful from a variational point of view for the total fractional-order variation to derive the first optimal conditions.

As discussed in Remark 2.1, the three definitions of fractional derivatives are equivalent under suitable conditions on continuity and boundary [16], and therefore from here on we shall drop the superscripts. Especially for fractional derivative operator defined on a full interval $[a, b]$, next, we shall denote them by D^α. Similarly for gradient and divergence operators, we denote by $\nabla^\alpha = (\frac{\partial^\alpha}{\partial x_1^\alpha}, \dots, \frac{\partial^\alpha}{\partial x_d^\alpha})^t$ and $\mathrm{div}^\alpha = \nabla^\alpha \cdot$. Thus $\frac{\partial^\alpha \phi}{\partial x_i^\alpha}$ denotes a fractional α-order derivative $D_{[a,b]}^\alpha \phi$ of ϕ along x_i direction.

3. A total fractional-order variation based image registration model

Our main aim below is to present a new registration model based on fractional differential information, which can trace the steady smooth transformation between two given images due to its nonlocal property, as illustrated in Fig. 2 which compares the TV and a fractional differential D^α restoration model (denoted by TV^α), and shows that the latter produced a better result. In the regularization framework (3), define $S(\boldsymbol{u})$ as a total fractional-order variation. We first introduce it in a rigorous and general setting, then consider its smoothing version and finally present our numerical algorithm.
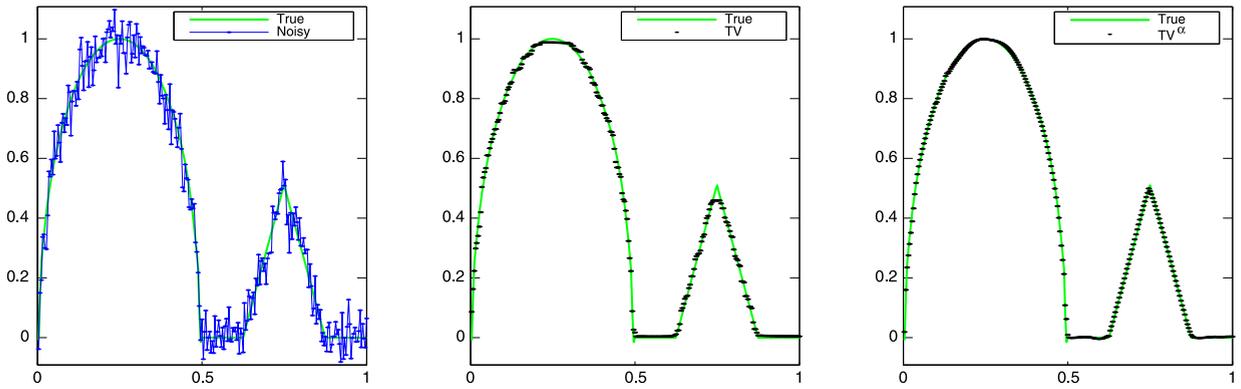
Fig. 2. Comparison of the TV regularizer [45] and the total fractional-order variation regularizer $\alpha = 1.6$ in restoring a noisy signal (far left).

3.1. The total fractional-order variation

In variational regularization methods, integration by parts involves the space of test functions in addition to the main solution space. Before discussing the total fractional-order variation, we give the following two definitions:

Definition 3.1 (*Spaces of test functions*). Denote by $\mathcal{C}^\ell(\Omega, \mathcal{R}^d)$ the space of an ℓ-order continuously differentiable functions in the image domain $\Omega \subset \mathcal{R}^d$. Then an ℓ-order compactly supported continuous function space as a subspace of $\mathscr{C}^\ell(\Omega, \mathcal{R}^d)$ is denoted by $\mathscr{C}_0^\ell(\Omega, \mathcal{R}^d)$, in which each member $v : \Omega \mapsto R^d$ satisfies the homogeneous boundary conditions $\frac{\partial^i v(x)}{\partial n^i}|_{\partial \Omega} = 0$ for all $i = 0, 1, \ldots, \ell$. Further, a special subspace of $\mathscr{C}_0^\ell(\Omega, \mathcal{R}^d)$ is denoted by

$$K := \left\{ \boldsymbol{\phi} \in \mathscr{C}_0^\ell(\Omega, \mathcal{R}^d) \,\middle|\, |\boldsymbol{\phi}(x)| \le 1 \text{ for all } x \in \Omega \right\}.$$

With a test function $\eta(x) \in \mathscr{C}_0^{n-1}(\Omega, \mathcal{R})$, the α-order integration by parts formulas can be rewritten as

$$\int_a^b \xi(x) \cdot D^\alpha \eta(x) dx = (-1)^n \int_a^b \eta(x) \cdot D^\alpha \xi(x) dx + \sum_{j=0}^{n-1} (-1)^j D^{\alpha-n+j} \xi(x) \frac{\partial^{n-j-1}\eta(x)}{\partial x^{n-j-1}} \Big|_{x=a}^{x=b}$$

$$= (-1)^n \int_a^b \eta(x) \cdot D^\alpha \xi(x) dx. \tag{5}$$

Definition 3.2 (*Total fractional-order variation*). For $0 \le \ell := n - 1 < \alpha < n$, define

$$S_\beta^\alpha(\varphi) := \sup_{\boldsymbol{\phi} \in K} \int_\Omega \left(-\varphi \, \mathrm{div}^\alpha \, \boldsymbol{\phi} + \sqrt{\beta(1 - |\boldsymbol{\phi}(x)|^2)} \right) dx, \tag{6}$$

as the general total fractional-order variation with a smoothing constraint with $\beta \ge 0$, where $\mathrm{div}^\alpha \boldsymbol{\phi} = \nabla^\alpha \cdot \boldsymbol{\phi}$, $|\boldsymbol{\phi}|^2 = \sum_{i=1}^d \phi_i^2$. If $\beta = 0$, $S_0^\alpha(\varphi)$ is called as the total fractional-order variation. Especially, if $1 < \alpha < 2$, i.e. $\ell = 1$ and $\varphi \in \mathcal{C}^\ell(\Omega, \mathcal{R}) \subseteq \mathcal{C}^1(\Omega, \mathcal{R})$, then $S_0^\alpha(\varphi) = \int_\Omega |\nabla^\alpha \varphi| dx$.

Lemma 3.1. *The total fractional-order variation $S_0^\alpha(\varphi)$ defines a semi-norm.*

Proof. Firstly we show that $S_0^\alpha(\varphi) \ge 0$. Note that $K = -K$ (i.e. $\boldsymbol{\phi} \in K$ if and only if $\xi = -\boldsymbol{\phi} \in K$); hence $\inf_{-\xi=\boldsymbol{\phi}\in K} \int_\Omega -\varphi \times \mathrm{div}^\alpha \xi dx = \inf_{\xi=-\boldsymbol{\phi}\in -K=K} \int_\Omega -\varphi \, \mathrm{div}^\alpha \, \xi dx = -\sup_{\boldsymbol{\phi}\in K} \int_\Omega -\varphi \, \mathrm{div}^\alpha \, \boldsymbol{\phi} dx$. Since $\sup_{\boldsymbol{\phi}\in K} \int_\Omega -\varphi \, \mathrm{div}^\alpha \, \boldsymbol{\phi} dx = \int_\Omega -\varphi \, \mathrm{div}^\alpha \, \boldsymbol{\phi}_0 dx$ implies $\inf_{\xi\in K} \int_\Omega -\varphi \, \mathrm{div}^\alpha \, \xi dx = \inf_{-\xi=\boldsymbol{\phi}\in K} \int_\Omega -\varphi \, \mathrm{div}^\alpha \, \xi dx = -\sup_{\boldsymbol{\phi}\in K} \int_\Omega -\varphi \, \mathrm{div}^\alpha \, \boldsymbol{\phi} dx = -\int_\Omega -\varphi \, \mathrm{div}^\alpha \, \boldsymbol{\phi}_0 dx$. If there exists a function φ such that $S_0^\alpha(\varphi) = \sup_{\boldsymbol{\phi}\in K} \int_\Omega -\varphi \, \mathrm{div}^\alpha \, \boldsymbol{\phi} dx < 0$, then $\inf_{\xi\in K} \int_\Omega -\varphi \, \mathrm{div}^\alpha \, \xi dx > 0$ which contradicts to the supremum bound. Therefore for any φ, $S_0^\alpha(\varphi) \ge 0$.

Secondly for any scalar c, we have $\sup_{\boldsymbol{\phi}\in K} \int_\Omega -(c\varphi) \, \mathrm{div}^\alpha(\boldsymbol{\phi}) dx = |c| \sup_{\mathrm{Sign}(c)\boldsymbol{\phi}\in K} \int_\Omega -\varphi \, \mathrm{div}^\alpha(\mathrm{Sign}(c)\boldsymbol{\phi}) dx$ since $\int_\Omega -(c\varphi) \times \mathrm{div}^\alpha(\boldsymbol{\phi}) dx = |c| \int_\Omega -\varphi \, \mathrm{div}^\alpha(\mathrm{Sign}(c)\boldsymbol{\phi}) dx$. Hence $S_0^\alpha(c\varphi) = |c| S_0^\alpha(\varphi)$.

Thirdly for any φ and ψ, $S_0^\alpha(\varphi + \psi) = \sup\limits_{\boldsymbol{\phi} \in K} \int_\Omega -(\varphi + \psi) \operatorname{div}^\alpha \boldsymbol{\phi} dx \leq \sup\limits_{\boldsymbol{\phi} \in K} \int_\Omega -\varphi \operatorname{div}^\alpha \boldsymbol{\phi} dx + \sup\limits_{\boldsymbol{\phi} \in K} \int_\Omega -\psi \operatorname{div}^\alpha \boldsymbol{\phi} dx = S_0^\alpha(\varphi) + S_0^\alpha(\psi)$. Thus we have proved that $S_0^\alpha(\varphi)$ is a semi-norm. $\square$

Actually for any $\alpha > 0$, the left R–L $D_{[0,x]}^\alpha f(x) = 0$ if $f(x) = x^{\alpha-k}(\neq 0)$ for all $k = 1, 2, \ldots, 1 + [\alpha]$ (with maximal integer part $[\alpha]$ such that $[\alpha] < \alpha$, refer to [46]), hence $S_0^\alpha(f) = \int_\Omega |\nabla^\alpha f| dx = 0$, which shows that $S_0^\alpha(\varphi)$ is not a norm. Equipped with the BV_2^α norm

$$\|\varphi\|_{BV_2^\alpha} = \|\varphi\|_{L^2} + S_0^\alpha(\varphi)$$

from Definition 3.2, the space of functions of α-bounded variation on Ω can be defined by

$$BV_2^\alpha(\Omega) := \left\{ \varphi \in L^2(\Omega) \mid S_0^\alpha(\varphi) < +\infty \right\}. \tag{7}$$

Lemma 3.2. *The space $BV_2^\alpha(\Omega)$ is a Banach space.*

Proof. Refer to [47] for more details of a related proof. $\square$

Alternatively, using the Riesz representation theorem [48,49], the total fractional-order variation of image $\varphi(x)$ may be regarded as the total variation of some positive Radon measure, i.e., bounded α-order variation (α-bounded variation) if it is bounded.

3.2. The proposed image registration model

We now introduce our proposed functional $E(\boldsymbol{u})$ in the optimization framework (3)

$$\min_{\boldsymbol{u} \in BV_2^\alpha(\Omega)} \left\{ E(\boldsymbol{u}) := S(\boldsymbol{u}) + \frac{\lambda}{2} D(\boldsymbol{u}) = \int_\Omega \psi\left(\sum_{i=1}^d |\nabla^\alpha u_i|\right) dx + \frac{\lambda}{2} \int_\Omega (T(x + \boldsymbol{u}(x)) - R(x))^2 dx \right\} \tag{8}$$

where $n - 1 < \alpha < n$ and $\boldsymbol{u}(x) = (u_1(x), \ldots, u_d(x))^t : \Omega \to R^d$. We shall mainly consider $d = 2$ though the procedure applies to a general d. Especially if taking $\psi(s) = s$, then one has $\psi(\sum_{i=1}^2 |\nabla^\alpha u_i|) = \sum_{i=1}^2 \|\nabla^\alpha u_i\|$. Since the objective $E(\boldsymbol{u})$ is not differentiable when $|\nabla^\alpha u_i| = 0$, it is common to consider a modified minimization problem of the form (8)

$$\min_{\boldsymbol{u} \in BV_2^\alpha(\Omega)} \left\{ E_\beta(\boldsymbol{u}) = \int_\Omega \psi\left(\sum_{i=1}^2 \sqrt{|\nabla^\alpha u_i|^2 + \beta}\right) dx + \frac{\lambda}{2} \int_\Omega (T(x + \boldsymbol{u}(x)) - R(x))^2 dx \right\}. \tag{9}$$

Discussion of the existence of $\boldsymbol{u}$ in (9) is complicated by the second (fidelity) term which is nonconvex. One simple idea is to replace $T(x + \boldsymbol{u}(x)) - R(x)$ by its linear approximation $T(x) - R(x) + \nabla T(x) \cdot \boldsymbol{u}(x)$; then one can use convexity to establish existence but the displacements cannot be too large. Here in this paper, we shall assume that the image functions (*Reference R* and *Template T*) are Lipschitz; hence functional $E_\beta(\boldsymbol{u})$ is coercive on $BV_2^\alpha(\Omega)$. Thus, we can uniformly bound the minimizing sequences and extract a converging subsequence for the $BV_2^\alpha - w^*$ topology which is defined as

$$u_j \xrightarrow[BV_2^\alpha - w^*]{*} u \quad \Longleftrightarrow \quad u_j \xrightarrow[L^2(\Omega)]{} u \quad \text{and} \quad \int_\Omega \boldsymbol{\phi} \cdot \nabla^\alpha u_j dx \to \int_\Omega \boldsymbol{\phi} \cdot \nabla^\alpha u dx$$

for all $\boldsymbol{\phi}$ in $\mathscr{C}_0^0(\Omega, \mathbf{R}^2)$. Since $E_\beta(\boldsymbol{u})$ is lower semi-continuous for this topology under Lipschitz function assumption, we can deduce the existence of a minimum [50,51].

For the registration model proposed, we should focus on deriving the Euler–Lagrange equation and considering its effective numerical solution methods.

Theorem 3.1. *Assume that $\psi(s) = s$ and $n - 1 < \alpha < n$; the Euler–Lagrange equation for problem (9) is the following*

$$(-1)^n \operatorname{div}^\alpha \left(\frac{\nabla^\alpha u_i(x)}{\sqrt{|\nabla^\alpha u_i(x)|^2 + \beta}} \right) + \lambda(T(x + \boldsymbol{u}(x)) - R(x)) \frac{\partial T(x + \boldsymbol{u}(x))}{\partial u_i} = 0, \text{ for all } i = 1, 2 \tag{10}$$

with one of these sets of boundary conditions for $n = 2$

1) *fixed boundary conditions* $u_i(x)\big|_{\partial\Omega} = b_1^i(x)$, *and* $\dfrac{\partial u_i(x)}{\partial n}\bigg|_{\partial\Omega} = b_2^i(x)$;

2) *homogeneous boundary conditions* $D^{\alpha-2} \dfrac{\nabla^\alpha u_i}{\sqrt{|\nabla^\alpha u_i|^2 + \beta}} \cdot \vec{n} = 0$, $\quad D^{\alpha-1} \dfrac{\nabla^\alpha u_i}{\sqrt{|\nabla^\alpha u_i|^2 + \beta}} \cdot \vec{n} = 0$

where $\boldsymbol{u} = \boldsymbol{u}(x)$ is the displacement field of the grid for template image $T = T(x)$, $x = (x_1, x_2) \in \Omega = (a, b) \times (c, d) \subset \mathbb{R}^2$ and $\vec{n}$ is the direction of the outward normal to the boundary $\partial\Omega$.

Note that the negative index for D in 2) implies that it is a fractional integral – see (4).

Proof. The proof can be done by using the first variation directly.

Firstly we define $\nabla E_\beta(\boldsymbol{u}) = \begin{pmatrix} \frac{\partial E_\beta(\boldsymbol{u})}{\partial u_1} \\ \frac{\partial E_\beta(\boldsymbol{u})}{\partial u_2} \end{pmatrix}$ and $\boldsymbol{\omega} = (\omega_1, \omega_2)^t$. To shorten the proof, let ω_i be a function in $C^1(\Omega)$ to be specified shortly for any $i = 1, 2$. For $u_i \in BV_2^\alpha(\Omega)$, we compute the first-order G-derivative (Gateaux) of the functional $E_\beta(u)$ in the direction ω_i by

$$\nabla E_\beta(\boldsymbol{u}) \cdot \boldsymbol{\omega} = \lim_{t \to 0} \frac{E_\beta(\boldsymbol{u} + t\boldsymbol{\omega}) - E_\beta(\boldsymbol{u})}{t} = \lim_{t \to 0} \frac{S_\beta^\alpha(\boldsymbol{u} + t\boldsymbol{\omega}) - S_\beta^\alpha(\boldsymbol{u})}{t} + \frac{\lambda}{2} \frac{D(\boldsymbol{u} + t\boldsymbol{\omega}) - D(\boldsymbol{u})}{t}. \tag{11}$$

Using the Taylor series w.r.t. t yields

$$S_\beta^\alpha(\boldsymbol{u} + t\boldsymbol{\omega}) - S_\beta^\alpha(\boldsymbol{u}) = \sum_{i=1}^2 \left(S_\beta^\alpha(u_i + t\omega_i) - S_\beta^\alpha(u_i) \right) = \sum_{i=1}^2 \int_\Omega \sqrt{|\nabla^\alpha u_i + t\nabla^\alpha \omega_i|^2 + \beta} - \sqrt{|\nabla^\alpha u_i|^2 + \beta} \, dx$$

and

$$\lim_{t \to 0} \frac{S_\beta^\alpha(u_i + t\omega_i) - S_\beta^\alpha(u_i)}{t} = \lim_{t \to 0} \frac{1}{t} \left[t \int_\Omega \boldsymbol{W}^i \cdot \nabla^\alpha \omega_i dx + O(t^2) \right] = \int_\Omega \boldsymbol{W}^i \cdot \nabla^\alpha \omega_i dx, \quad \boldsymbol{W}^i = \frac{\nabla^\alpha u_i}{\sqrt{|\nabla^\alpha u_i|^2 + \beta}}. \tag{12}$$

The quadratic L_2 data fidelity term is easy to deal with

$$\lim_{t \to 0} \frac{D(\boldsymbol{u} + t\boldsymbol{\omega}) - D(\boldsymbol{u})}{t} = \lim_{t \to 0} \int_\Omega \frac{(T(x + \boldsymbol{u} + t\boldsymbol{\omega}) - R(x))^2 - (T(x + \boldsymbol{u}) - R(x))^2}{t} dx$$

$$= 2 \int_\Omega (T(x + \boldsymbol{u}) - R) \nabla_{\boldsymbol{u}} T(x + \boldsymbol{u}) \cdot \boldsymbol{\omega} \, dx. \tag{13}$$

Recall that

$$\int_\Omega \boldsymbol{W}^i \cdot \nabla^\alpha \omega_i dx = (-1)^n \int_\Omega \omega_i \, \text{div}^\alpha \, \boldsymbol{W}^i dx + \sum_{j=0}^{n-1} (-1)^{n+j} \int_c^d D_{[a,b]}^{\alpha-n+j} W_1^i \frac{\partial^{n-j-1} \omega_i(x)}{\partial x_1^{n-j-1}} \Big|_{x_1=a}^{x_1=b} dx_2$$

$$+ \sum_{j=0}^{n-1} (-1)^{n+j} \int_a^b D_{[c,d]}^{\alpha-n+j} W_2^i \frac{\partial^{n-j-1} \omega_i(x)}{\partial x_2^{n-j-1}} \Big|_{x_2=c}^{x_2=d} dx_1 \tag{14}$$

where we note $n = 2$ for $1 < \alpha < 2$. Next consider 2 cases:

1). Given $u_i(x)\big|_{\partial\Omega} = b_1^i(x)$, and $\frac{\partial u_i(x)}{\partial n}\big|_{\partial\Omega} = b_2^i(x)$, since $(u_i(x) + t\omega_i(x))\big|_{\partial\Omega} = (u_i(x))\big|_{\partial\Omega} = b_1^i(x)$ and $\frac{\partial(u_i(x)+t\omega_i(x))}{\partial n}\big|_{\partial\Omega} = \frac{\partial u(x)}{\partial n}\big|_{\partial\Omega} = b_2^i(x)$, it suffices to take $\omega_i \in \mathscr{C}_0^1(\Omega, \mathcal{R})$. From Lemma 2.1, such a choice ensures $\frac{\partial^k \omega(x)}{\partial n^k}\big|_{\partial\Omega} = 0, k = 0, 1 \Rightarrow \frac{\partial^{n-j-1}\omega_i(x)}{\partial x_1^{n-j-1}}\big|_{x_1=a \text{ or } b} = \frac{\partial^{n-j-1}\omega_i(x)}{\partial x_1^{n-j-1}}\big|_{x_2=c \text{ or } d} = 0, n - j - 1 = 0, 1$. Hence Eq. (11) with (12) reduces to (10).

2). Keep $\omega_i \in C^1(\Omega)$. Since $\frac{\partial^{n-j-1}\omega_i(x)}{\partial x_1^{n-j-1}}\big|_{x_1=a \text{ or } b} \neq 0$, $\frac{\partial^{n-j-1}\omega_i(x)}{\partial x_1^{n-j-1}}\big|_{x_2=c \text{ or } d} \neq 0$, the boundary terms in Eq. (14) can only diminish if

$$D_{[a,b]}^{\alpha-n+j} W_1^i \Big|_{x_1=a \text{ or } b} = 0 \quad \text{and} \quad D_{[c,d]}^{\alpha-n+j} W_2^i \Big|_{x_2=c \text{ or } d} = 0 \quad \Rightarrow \quad D^{\alpha-n+j} \boldsymbol{W}^i \cdot \vec{n} = 0, \ j = 0, 1.$$

The proof is complete. □

Remark 3.1 (*Boundary condition*). In non-rigid image deformation applications, the displacement field $\boldsymbol{u}(x)$ of image coordination can be required as an ℓ-compact support continuous function in Ω. Therefore the above first set 1) of boundary conditions seems simple and reasonable, because one easily knows a priori what $\boldsymbol{u}(x)$ on boundary $\partial\Omega$ should be (i.e. $u_i(x)\big|_{\partial\Omega} = b_1^i(x) = 0$ and $\frac{\partial u_i(x)}{\partial n}\big|_{\partial\Omega} = b_2^i(x) = 0$ are known from an application view). In fact, these boundary and continuity conditions are also used for the equivalence of three definitions (RL, GL, C).

Remark 3.2. The second set 2) of boundary conditions appears complicated which might be simplified as [16, Section 2.3.6, p. 75], [47]:

$$\frac{\partial^k u_i(x)}{\partial x_j^k}\bigg|_{\partial\Omega} = 0 \quad \text{for all} \quad i = 1, 2; \; j = 1, 2; \; k = 0, 1, 2.$$

As we have seen in [16, Section 2.3.6, p. 75], if $W_1^i(x)$ has a sufficient number of continuous derivatives, then $D_{[a,b]}^{\alpha-n+j} W_1^i\big|_{x_1 = a \text{ or } b} = 0$ is equivalent to $\frac{\partial^j W_1^i}{\partial x_1^j}\big|_{x_1 = a \text{ or } b} = 0 \; (j = 0, 1)$, i.e.,

$$W_1^i\bigg|_{x_1 = a \text{ or } b} = 0 \quad \text{and} \quad \frac{\partial W_1^i}{\partial x_1}\bigg|_{x_1 = a \text{ or } b} = 0.$$

Indeed, if n-th derivative of $u_i(x)$ is integrable in $[a, b]$, then $W_1^i\big|_{x_1 = a \text{ or } b} = 0$ is equivalent to the conditions

$$u_i(x)\bigg|_{x_1 = a \text{ or } b} = 0 \quad \text{and} \quad \frac{\partial u_i(x)}{\partial x_1}\bigg|_{x_1 = a \text{ or } b} = 0,$$

on the other hand, $\frac{\partial^k u_i(x)}{\partial x_1^k}\big|_{x_1 = a \text{ or } b} = 0$ (for all $k = 0, 1, 2$) are equivalent to $\frac{\partial^\alpha u_i(x)}{\partial x_1^\alpha}\big|_{x_1 = a \text{ or } b} = 0$ and $\frac{\partial^{1+\alpha} u_i(x)}{\partial x_1^{1+\alpha}}\big|_{x_1 = a \text{ or } b} = 0$, hence one has $\frac{\partial W_1^i}{\partial x_1}\big|_{x_1 = a \text{ or } b} = 0$. The derivations of W_2^i are similar to those of W_1^i.

3.3. The solution algorithm

Effective numerical implementation is of great importance. Next we present our iterative methods for solving the resulting nonlinear system with structured block matrices, following a finite difference discretization. Of course, for nonlinear systems, there exist many alternative ways of iterative solution that one might consider or develop.

3.3.1. A semi-implicit scheme

To solve (10) for $1 < \alpha < 2$, we just update $\boldsymbol{u}$ to the steady state form of the following fractional-order evolution equation with time:

$$\frac{\partial u_i}{\partial t} = -\text{div}^\alpha\left(\frac{\nabla^\alpha u_i(x)}{\sqrt{|\nabla^\alpha u_i(x)|^2 + \beta}}\right) - \lambda(T(x + \boldsymbol{u}) - R(x))\frac{\partial T(x + \boldsymbol{u})}{\partial u_i}, \quad \text{for } i = 1, 2 \tag{15}$$

with the ℓ-compact support boundary conditions (combination of zero-Neumann and zero-Dirichlet conditions, *see* Remark 3.1), which may be discretized semi-implicitly along time step once an initial guess u_0 is given, i.e.,

$$\frac{u_i^{n+1} - u_i^n}{\Delta t} = -\text{div}^\alpha\left(\frac{\nabla^\alpha u_i^{n+1}(x)}{\sqrt{|\nabla^\alpha u_i^n(x)|^2 + \beta}}\right) - \lambda(T(x + \boldsymbol{u}^n) - R(x))\frac{\partial T(x + \boldsymbol{u}^n)}{\partial u_i},$$

which yields the following linear systems

$$\left(I + \Delta t\, \text{div}^\alpha\left(\frac{\nabla^\alpha}{\sqrt{|\nabla^\alpha u_i^n(x)|^2 + \beta}}\right)\right) u_i^{n+1}(x) = u_i^n - \lambda\Delta t(T(x + \boldsymbol{u}^n) - R(x))\frac{\partial T(x + \boldsymbol{u}^n)}{\partial u_i}. \tag{16}$$

3.3.2. Spacial discretization

Before introducing the discretization of the fractional derivative, we define a spatial partition (x_k, y_l) (for all $k = 0, 1, \ldots, N + 1; l = 0, 1, \ldots, M + 1$) of image domain Ω. Here we mainly consider the discretization of the α-order fractional derivative at the inner point (x_k, y_l) (for all $k = 0, 1, \ldots, N; l = 0, 1, \ldots, M$) on Ω along x-direction by using the shifted Grünwald approximation approach as [16,20]

$$D_{[a,b]}^\alpha f(x_k, y_l) = \frac{\delta_0^\alpha f(x_k, y_l)}{h^\alpha} + O(h) = \frac{1}{2}\left(\frac{\delta_-^\alpha f(x_k, y_l)}{h^\alpha} + \frac{\delta_+^\alpha f(x_k, y_l)}{h^\alpha}\right) + O(h)$$

$$= \frac{1}{2}\left(h^{-\alpha}\sum_{j=0}^{k+1} \rho_j^\alpha f_{k-j+1}^l + h^{-\alpha}\sum_{j=0}^{N-k+2} \rho_j^\alpha f_{k+j-1}^l\right) + O(h), \tag{17}$$

where $f_s^l := f_{s,l}$, $\rho_j^{(\alpha)} = (-1)^j \binom{\alpha}{j}$, $j = 0, 1, \ldots, N$ and $\rho_0^{(\alpha)} = 1$; $\rho_j^{(\alpha)} = (1 - \frac{1+\alpha}{j})\rho_{j-1}^{(\alpha)}$, for $j > 0$. Discretization of fractional derivatives in the Fourier space is presented in [23,24].

As we work in 2D with $d = 2$, it is easy to see from (17) that the first order estimate of the α-order fractional $D^{\alpha}_{[a,b]} f(x_k, y_l)$ along x-direction at the point (x_k, y_l) with a fixed y_l is a linear combination of $N + 2$ values $\{f^l_0, f^l_1, \ldots, f^l_N, f^l_{N+1}\}$. One common approach dealing with discrete approximation is based on periodic boundary condition in such a way that an image defined on Ω will be extended symmetrically about its borders [23]. Here incorporating the integration by parts formulas and the equivalence of three fractional derivative definitions (*see* also Remark 2.1), a zero boundary condition on $\partial\Omega$ is considered into matrix approximation of fractional derivative (as done in [20,52]); hence all N equations of fractional derivatives along x direction in formulas (17) can be written simultaneously in the matrix form:

$$
\begin{pmatrix} \delta^{\alpha}_0 f(x_1, y_l) \\ \delta^{\alpha}_0 f(x_2, y_l) \\ \vdots \\ \vdots \\ \delta^{\alpha}_0 f(x_N, y_l) \end{pmatrix} = \frac{1}{2} \underbrace{\begin{pmatrix} 2\rho^{\alpha}_1 & \rho^{\alpha}_0 + \rho^{\alpha}_2 & \rho^{\alpha}_3 & \cdots & & \rho^{\alpha}_N \\ \rho^{\alpha}_0 + \rho^{\alpha}_2 & 2\rho^{\alpha}_1 & \ddots & \ddots & & \vdots \\ \rho^{\alpha}_3 & \ddots & \ddots & \ddots & & \rho^{\alpha}_3 \\ \vdots & \ddots & \ddots & 2\rho^{\alpha}_1 & \rho^{\alpha}_0 + \rho^{\alpha}_2 \\ \rho^{\alpha}_N & \cdots & \rho^{\alpha}_3 & \rho^{\alpha}_0 + \rho^{\alpha}_2 & 2\rho^{\alpha}_1 \end{pmatrix}}_{B_N} \underbrace{\begin{pmatrix} f^l_1 \\ f^l_2 \\ \vdots \\ \vdots \\ f^l_N \end{pmatrix}}_{f}. \tag{18}
$$

From the approximation of fractional order derivative (17), for any $1 < \alpha < 2$, the coefficients $\rho^{(\alpha)}_k$ suffice to show the following properties [16,18]:

$$
1)\ \rho^{(\alpha)}_0 = 1,\ \rho^{(\alpha)}_1 = -\alpha < 0,\ 1 \ge \rho^{(\alpha)}_2 \ge \rho^{(\alpha)}_3 \ge \ldots \ge 0;\quad 2)\ \sum_{k=0}^{\infty} \rho^{(\alpha)}_k = 0,\ \sum_{k=0}^{m} \rho^{(\alpha)}_k \le 0\ (m \ge 1).
$$

Hence by the Gerschgorin circle theorem, one can deduce that matrix B_N in (18) is a symmetric and negative definite Toeplitz matrix (i.e., $-B_N$ is a positive definite Toeplitz matrix).

Consider the nodes $(kh_x;\ lh_y)$, $k = 0, 1, \ldots, N + 1$, corresponding to x-direction spatial discretization nodes for a y-direction $l = 0, 1, \ldots, M + 1$, i.e. sorted in lexico-graphical order. So all values of α-th order x-direction derivative of $\upsilon(x; y)$ at these nodes are approximated using the discrete analogue of differentiation of arbitrary α order derivative:

$$
\upsilon^{(\alpha)}_x = (I_M \otimes B_N)\upsilon = B^{(\alpha)}_x \upsilon,
$$

where

$$
\upsilon^{(\alpha)}_x = \left(\upsilon^{(\alpha)}_{11}, \ldots, \upsilon^{(\alpha)}_{N1}, \upsilon^{(\alpha)}_{12}, \ldots, \upsilon^{(\alpha)}_{NM} \right)^T,\ \upsilon = (\upsilon_{11}, \ldots, \upsilon_{N1}, u_{12}, \ldots, \upsilon_{NM})^T,
$$

and $\otimes$ denotes the Kronecker product, for example, $A \otimes B = [a_{ij}B]$ is the $np \times mq$ matrix for the $p \times q$ matrix A and $n \times m$ matrix B. Similarly, all values of α-th order y-direction derivative of $\upsilon(x; y)$ at these nodes are approximated by:

$$
\upsilon^{(\alpha)}_y = B^{(\alpha)}_y \upsilon = (B_M \otimes I_N)\upsilon,\ \text{ where } \upsilon^{(\alpha)}_y = \left(\upsilon^{(\alpha)}_{11}, \ldots, \upsilon^{(\alpha)}_{1M}, \upsilon^{(\alpha)}_{21}, \ldots, \upsilon^{(\alpha)}_{NM} \right)^T.
$$

Until now, we have described enough mathematical tools and notation necessary for approximating a fractional order derivative in variational image inverse problems using finite differences. To summarize, the discrete scheme of the fractional Euler–Lagrange equation (10) is then given by

$$
\left(u^{n+1}_i + \Delta t (B^{(\alpha)}_x)^T \left(\mathcal{HI}(|B^{(\alpha)}_x u^n_i|^{\epsilon}_H) \circ (B^{(\alpha)}_x u^{n+1}_i) \right) + (B^{(\alpha)}_y)^T \left(\mathcal{HI}(|B^{(\alpha)}_y u^n_i|^{\epsilon}_H) \circ (B^{(\alpha)}_y u^{n+1}_i) \right) \right)
$$

$$
= u^n_i - \lambda \Delta t (T(x + \mathbf{u}^n) - R(x)) \frac{\partial T(x + \mathbf{u}^n)}{\partial u_i},
$$

where $(\mathcal{HI}(A))_{ij} = A^{-1}_{ij}$ means a Hadamard inverse of a non-zero matrix A. The above represents a linear system

$$
G(u^n_i)u^{n+1}_i := \left(I + \Delta t \left((B^{(\alpha)}_x)^T \text{Diag}(\mathcal{HI}(|B^{(\alpha)}_x u^n_i|^{\epsilon}_H)) B^{(\alpha)}_x + (B^{(\alpha)}_y)^T \text{Diag}(\mathcal{HI}(|B^{(\alpha)}_y u^n_i|^{\epsilon}_H)) B^{(\alpha)}_y \right) \right) u^{n+1}_i
$$

$$
= \underbrace{u^n_i - \lambda \Delta t (T(x + \mathbf{u}^n) - R(x)) \frac{\partial T(x + \mathbf{u}^n)}{\partial u_i}}_{f^n_i}, \tag{19}
$$

with $\text{Diag}(A)$ the diagonal matrix of A.

3.3.3. The overall algorithm

A semi-implicit scheme for a gradient descent method for Eqs. (15) is implemented in our numerical experiments. For motivation purposes of discretization of the Euler–Lagrange equation, we are going to utilize many extensive properties of the Kronecker product. Firstly, if I is an identity matrix, then one has that $(I \otimes B)^T = I \otimes B^T$, $(A \otimes I)^T = A^T \otimes I$. Further, if x is an mq vector, the $m \times q$ matrix X is the reshaped vector x along its column, one has

$$[(A \otimes B)x]_{(i-1)m+j} = \sum_{l=1}^{q} \sum_{k=1}^{m} A_{i,l} B_{j,k} X_{k,l} = \sum_{l=1}^{q} \Big(\sum_{k=1}^{m} B_{j,k} X_{k,l} \Big) (A^T)_{l,i} = (BXA^T)_{j,i},$$

hence multiplication of the matrix $A \otimes B$ to a vector x can be by the matrix scheme BXA^T (i.e., $[(A \otimes B)x]_s = [BXA^T]_{j,i}$ with $s = (i-1)m + j$). Therefore, α-th order derivative $\upsilon_x^{(\alpha)}$ of $\upsilon(x; y)$ along x-direction at all nodes on Ω can be given by matrix $B_N U$, and similarly $U B_M^T$ for y-direction.

We consider a matrix approximation to the discrete semi-implicit scheme (19) evolving the Euler–Lagrange equation of the fractional regularization system with $1 < \alpha < 2$ as follows:

$$\underbrace{U^{n+1} + \Delta t \Big(B_N^T \big(\mathcal{H}\mathcal{I}(|B_N U^n|_H^\epsilon) \circ (B_N U^{n+1}) \big) + \big(\mathcal{H}\mathcal{I}(|U^n B_M^T|_H^\epsilon) \circ (U^{n+1} B_M^T) \big) B_M \Big)}_{A(U^n)U^{n+1}} = F_i^n,$$

where U is an $N \times M$-size reshaped matrix of vector u_i and F_i^n is an $N \times M$-size reshaped matrix of vector f_i^n, B_N and B_M are symmetric Toeplitz matrices as shown in Eq. (18).

To apply a CG method to Eq. (19), the next theorem guarantees the feasibility of our algorithms.

Theorem 3.2. *For any $NM \times 1$ vector φ^0 and $\Delta t > 0$, $G(\varphi^0)$ is a symmetric and positive definite matrix.*

Proof. From Eq. (19), for any non-zero vectors ϕ and ψ, we have easily

$$\phi^T G(\varphi^0)\psi = \langle \phi, G(\varphi^0)\psi \rangle$$
$$= \Delta t \, \phi^T \Big((B_x^{(\alpha)})^T \mathrm{Diag}(\mathcal{H}\mathcal{I}(|B_x^{(\alpha)}\varphi^0|_H^\epsilon)) B_x^{(\alpha)} + (B_y^{(\alpha)})^T \mathrm{Diag}(\mathcal{H}\mathcal{I}(|B_y^{(\alpha)}\varphi^0|_H^\epsilon)) B_y^{(\alpha)} \Big) \psi + \langle \phi, \psi \rangle$$
$$= \Delta t \Big(\langle B_x^{(\alpha)}\phi, B_x^{(\alpha)}\psi \rangle_{D_1} + \langle B_y^{(\alpha)}\phi, B_y^{(\alpha)}\psi \rangle_{D_2} \Big) + \langle \phi, \psi \rangle$$

with weighted positive matrices $D_1 = \mathrm{Diag}(\mathcal{H}\mathcal{I}(|B_x^{(\alpha)}\varphi^0|_H^\epsilon))$ and $D_2 = \mathrm{Diag}(\mathcal{H}\mathcal{I}(|B_y^{(\alpha)}\varphi^0|_H^\epsilon))$. Especially, taking $\psi = \phi$, one can show

$$\phi^T G(\varphi^0)\phi = \Delta t(\|B_x^{(\alpha)}\phi\|_{D_1}^2 + \|B_y^{(\alpha)}\phi\|_{D_2}^2) + \|\phi\|^2 > 0,$$

according to the definition of positive definite matrices, which proves that $G(\varphi^0)$ is a positive definite matrix. On the other hand, the above equation indicated that $\phi^T G(\varphi^0)\psi = \psi^T G(\varphi^0)\phi$, which shows that $G(\varphi^0)$ is a symmetric matrix. □

Remark 3.3. In the above proof, simply $B_x^{(\alpha)}u$, $B_y^{(\alpha)}u$ and u are replaced by $B_N U$, $U B_M^T$ and U respectively, then we show easily that the matrices product generating the vectors inner product is positive, i.e., $G(u^0)$ is a positive definite matrix if and only if the weighted matrices inner product $\langle WU, U \rangle = \sum_{ij}(\sum_k W_{ik}U_{kj})U_{ij}$ is positive, where $W = A(U^0)$.

A pseudo-code implementing the fixed point CG based iterative scheme is summarized in Algorithm 3.1:

Algorithm 3.1 *(The total fractional-order variation image registration).*

1.) Given initial values $U_i^0 = 0$ for $i = 1, 2$;
2.) Solve the linear system $A(U_i^n)U_i = F_i^n$ to obtain the unique solution U_i^{n+1} by using the CG algorithm, where we supply the multiplication $Y = A(U_i^n)U_i$ implicitly for the algorithm;
3.) Check the stopping condition: $\max_i \frac{|U_i^n - U_i^{n+1}|}{|U_i^n|} < \epsilon$,
   - If satisfied, stop and return $U_i^* := U_i^{n+1}$; then update the registered image by image interpolation;
   - else   set $n := n + 1$ and return to 2).

For practical implementations, our algorithm is also stopped if the maximum number of evolving iterations is reached (usually $it = 10\,000$), or the relative error $\max_i \frac{|U_i^n - U_i^{n+1}|}{|U_i^n|}$ in two consecutive iterative steps is smaller than a small number $\epsilon > 0$ (typically $\epsilon = 10^{-10}$ for a practical registration).

Table 1
Comparisons for the regularization parameter λ-independence, where we fix $\alpha = 1.6$.

$\lambda = e^{-4}$	0.01	0.1	0.5	1.0	1.5	**2.0**	2.5	3.0	3.5	4.0	5.0	10.0	30.0
SSD(e^6)	74.52	3.76	1.31	0.94	0.80	**0.78**	1.29	1.85	2.29	2.57	2.96	3.48	4.47
Re-SSD (%)	52.45	2.65	0.92	0.66	0.56	**0.55**	0.91	1.30	1.61	1.81	2.09	2.45	3.15

Table 2
Comparisons for the fractional order α-independence.

α	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9
SSD(e^6)	1.04	0.95	0.89	0.84	0.81	0.78	0.75	0.72	0.70
Re-SSD (%)	0.73	0.67	0.63	0.59	0.57	0.55	0.53	0.50	0.49

4. Numerical experiments

To illustrate the performance of our new model, some numerical results on several test problems are presented. Here we first test the effectiveness and integrity of the proposed algorithm and then present comparisons with leading models for registration. To assess $\phi(\boldsymbol{u}(x)) = x + \boldsymbol{u}(x)$, we compare

- the solution's visual quality (including the errors between the registered image and reference image);
- the *SSD* (the Sum of Squared Differences) and the *Re-SSD* (the relative Sum of Squared Differences) values which are given by [5,8]

$$SSD(T, R) = \frac{1}{2} \sum_{i,j} (T_{i,j} - R_{i,j})^2 \quad \text{and} \quad Re\text{-}SSD(T, R, T_{reg}) = \frac{SSD(T_{reg}, R)}{SSD(T, R)};$$

- *MFN* (the mesh folding number) and *MFR* (the mesh folding ratio) values which are given as

$$MFN(\phi(\boldsymbol{u})) = \sharp\Big(\det(J(\phi(\boldsymbol{u}(x)))) \leq 0 \Big) \quad \text{and} \quad MFR(\phi(\boldsymbol{u})) = \frac{MFN(\phi(\boldsymbol{u}))}{\sharp(\Omega^h)} \tag{20}$$

where T_{reg} is the registered image of the template image T, $\det(J(\phi(\boldsymbol{u}(x))))$ denotes the determinant of the Jacobian $J(\phi) = \begin{vmatrix} \frac{\partial\phi_1(x)}{\partial x_1} & \frac{\partial\phi_1(x)}{\partial x_2} \\ \frac{\partial\phi_2(x)}{\partial x_1} & \frac{\partial\phi_2(x)}{\partial x_2} \end{vmatrix}$ at x and $\sharp(\Omega^h)$ denotes the number of nodes in the discretized grid Ω^h.

It should be noted however that these valuation levels not always correlate with human perception, in real life situation, such measures should be interpreted with some caution because the true solution is not known.

4.1. Tests for parameters α- and λ-dependence

The purpose of this test set of examples is to show how sensitive our total fractional-order variation image registration model is with respect to values of differential order α and regularization parameter λ. On one hand, in image inverse problems a regularization parameter λ balancing the trade-off between a good fit to the data and a regular solution is difficult to fix: if the value λ is too large, then the corresponding displacement field $\boldsymbol{u}(x)$ is not one to one, while if too small, it is a poor transformation for matching between reference and template images. On the other hand, the fractional derivatives are defined using integrals which depend on the values of the function over the entire range of integration, so they are non-local. Although in the recent years this non-local property is very useful and popular for constructing simple material models and unified principles, it is a challenge to develop effective numerical computations and different α has been devoted to different combination of the values of the function over the entire range of integration, hence leading to different regularization properties of displacement field $\boldsymbol{u}(x)$.

Here we first assess how our model is affected when varying λ. To this end, Algorithm 3.1 was tested for a synthetic image with the results shown in Table 1, where a 256×256 piecewise constant circle image of range $[0, 255]$ is needed to register into a piecewise constant box image. Here λ is varied from $0.01e^{-4}$ to $30e^{-4}$. The selection of suitable λ is a separate but important issue because it is in general unknown a priori and it significantly affects on the qualities of registered images as well as the algorithm performance [53]. However, for the range of tested in Table 1, the proposed total fractional-order variation regularization model still obtains the satisfactory solution in a reasonable range of values λ, so for this example, the accurate selection of λ is not needed as any λ between $0.1e^{-4}$ and $10e^{-4}$ can give better results, and is reasonable and recommendable. The relationships of SSD/Re-SSD and parameter λ are shown in Fig. 3(a).

Next we test how our model with the total fractional-order variation regularization is affected for the above example when varying the values of α; α is varied from 1.1 to 1.9, Table 2 shows that our model solves the total fractional-order variation regularization image denoising inverse problem. As already discussed previously, the smaller α will lead to much
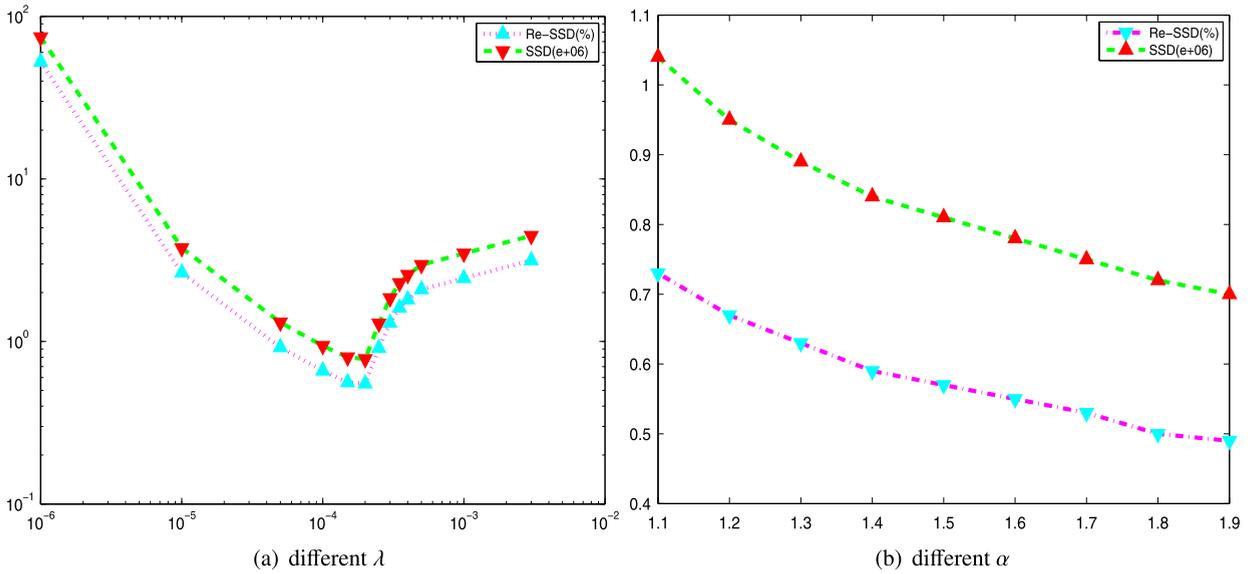
(a) different λ (b) different α

Fig. 3. Test 1 – comparisons of SSD and Re-SSD for different parameters λ and α.

(a) reference R   (b) template T   (c) Registered T_frac

(d) T-R   (e) T_frac-R   (f) transformation   (g) displacement field

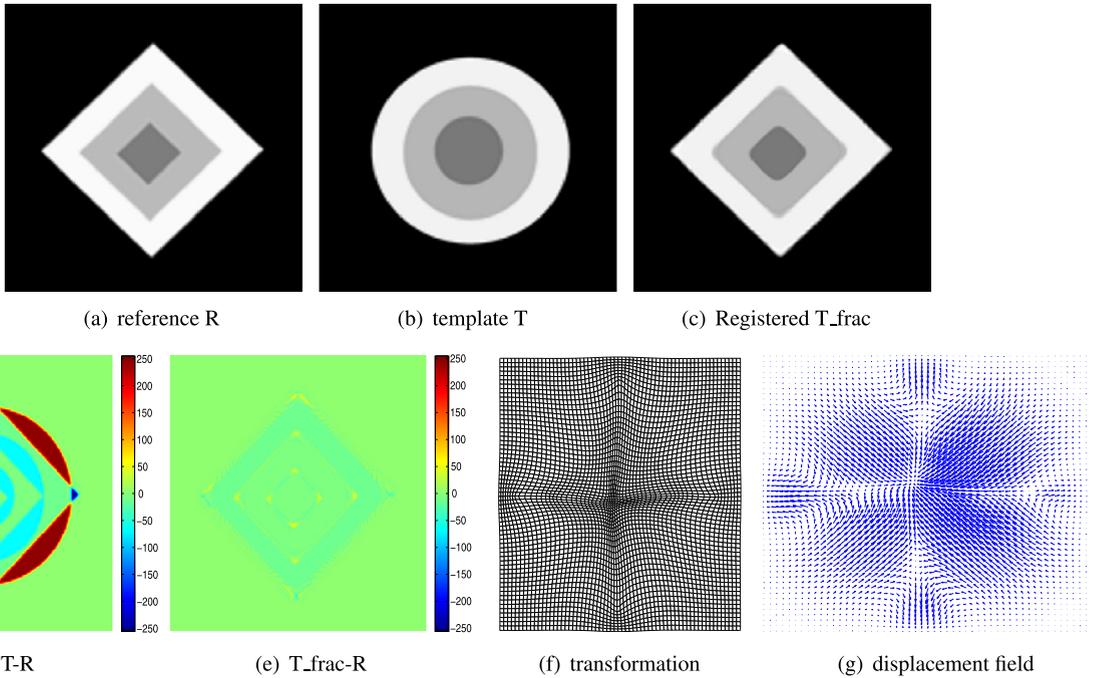**Fig. 4.** Test 1 – results (Re_SSD = 0.55%, *MFN* = 0 and *MFR* = 0%) of the example 1, where $\lambda = 2.0e^{-4}$ and $\alpha = 1.8$.

more blocky (staircase) effects of the displacement field and the larger α will make solution $\boldsymbol{u}(x)$ more smoother. Their dependent relationship is also shown in Fig. 3(b).

Further, for a fixed $\lambda = 2.0e^{-4}$ and $\alpha = 1.8$, the solution results by the proposed Algorithm 3.1 based on CG fixed update iterations are shown in Fig. 4 with the reference image and template image shown in Fig. 4(a) and Fig. 4(b), a registered image from T to R is presented in Fig. 4(c). The dissimilarity between T and R shown in Fig. 4(d) is reduced into Fig. 4(e), and the transformation of coordinate grid and displacement field $\boldsymbol{u}(x)$ are displayed in Fig. 4(f) and Fig. 4(g) respectively.

4.2. Tests for different pairs of fonts of alphabet

In order to exhibit the performance of our algorithm, a synthetic example of Alphabet registration problem from a font (*Aparajit* font) to another font (*Bodoni* font) shown in the first and second rows of Fig. 5 is first implemented. Through
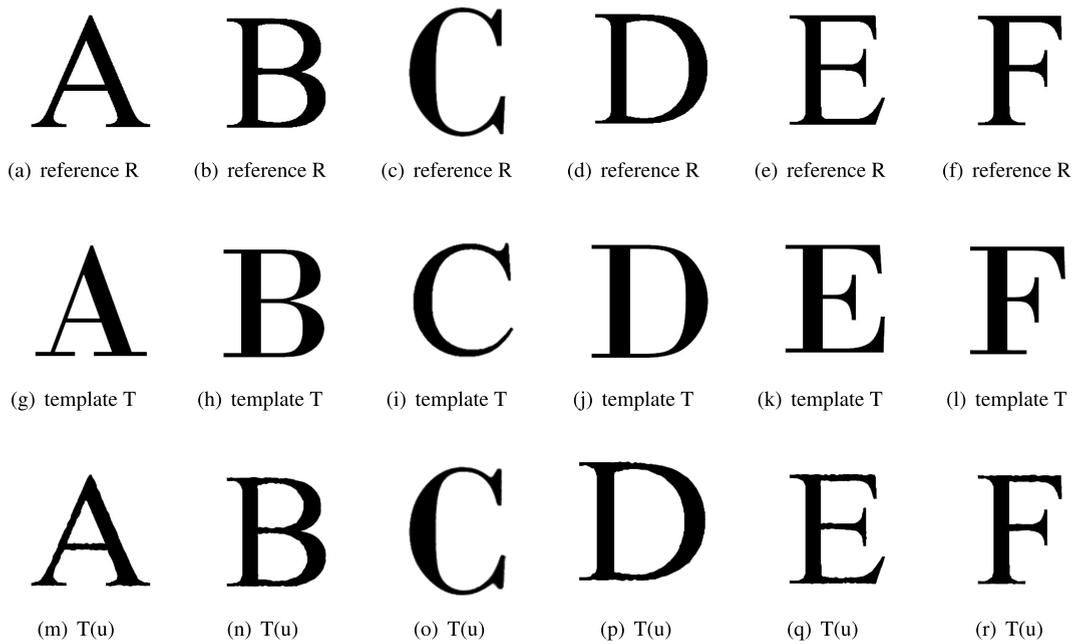
(a) reference R   (b) reference R   (c) reference R   (d) reference R   (e) reference R   (f) reference R

(g) template T   (h) template T   (i) template T   (j) template T   (k) template T   (l) template T

(m) T(u)   (n) T(u)   (o) T(u)   (p) T(u)   (q) T(u)   (r) T(u)

Fig. 5. Test 2 – presentations of registration for different fonts of alphabet.

(a) A ($MFN=0,MFR=0\%$)   (b) B ($MFN=0,MFR=0\%$)   (c) C ($MFN=0,MFR=0\%$)

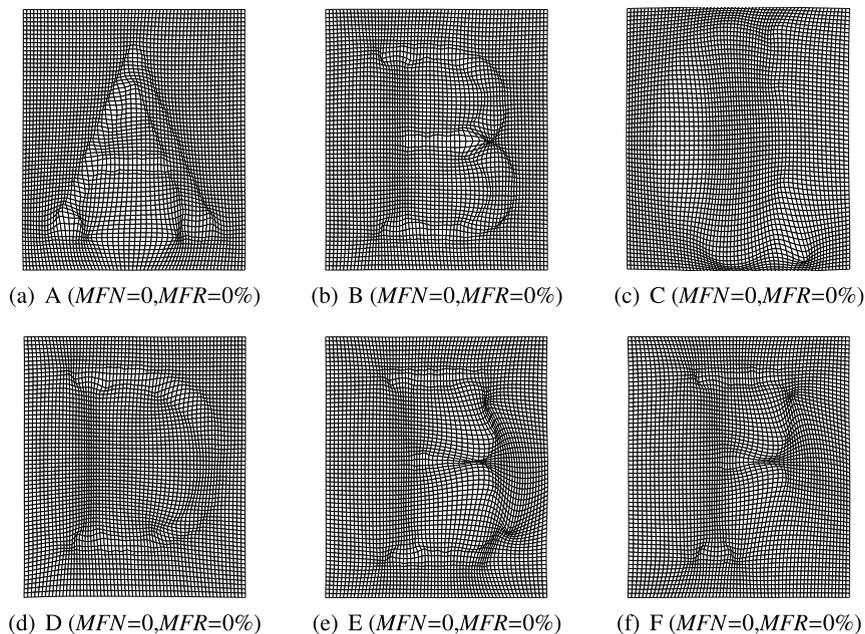(d) D ($MFN=0,MFR=0\%$)   (e) E ($MFN=0,MFR=0\%$)   (f) F ($MFN=0,MFR=0\%$)

Fig. 6. Test 2 – transformation presentations for different fonts of alphabet (mesh plots).

employing the fractional-order total variation image registration algorithm to register the two groups of Alphabet images where the second image is warped so that it is indistinguishable from the first one, the deformation minimizing the energy functional implies a correspondence between the two shapes. As the deformation is calculated in the form of a deformation field, we can display the transformation of the points on the first shape and see to which point on the other shape they are deformed to, and the registered shapes shown in the third row in Fig. 5 and reference shapes are matched perfectly.

As is shown in Fig. 6, apart from a small influence from the smoothing of the deformation field, the results are in fact the most plausible transformations to match those two groups of shapes.

Finally, we have also showed that the arrows in Fig. 7 indeed go exactly from one shape to the other, which is equivalent to an exact match.
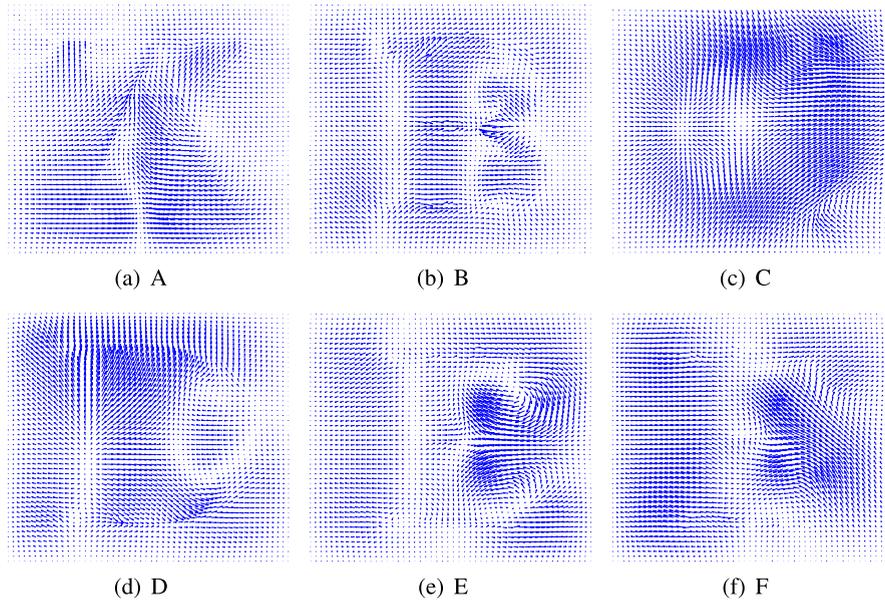
(a) A   (b) B   (c) C

(d) D   (e) E   (f) F

Fig. 7. Test 2 – displacement fields for different fonts of alphabet (flow maps).

(a) Reference R   (b) Template T   (c) diffusion (Re_SSD=20.44%)   (d) our fractional (Re_SSD=9.65%)
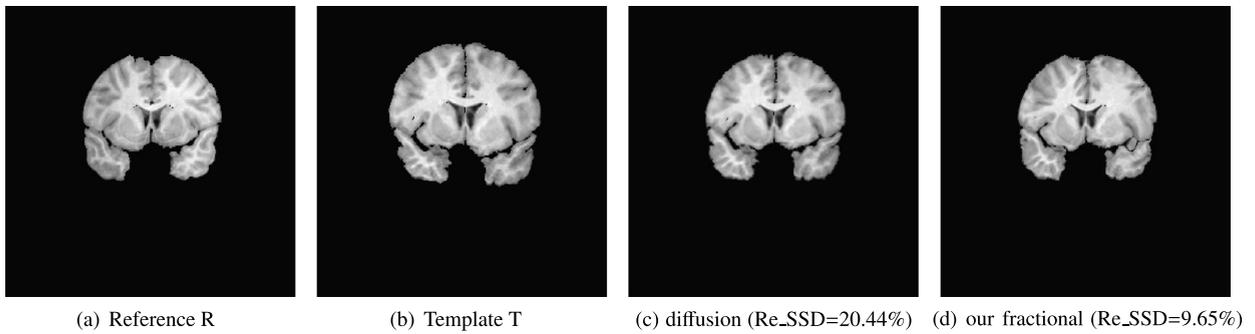
Fig. 8. Result presentations – comparisons between diffusion and our fractional.

4.3. Comparison with other variational models

It is of interest to compare our Algorithm 3.1 with other variational models in Section 2.3. Here we compare with the diffusion model in Figs. 8–9 and with the elastic, linear curvature and the mean curvature models in Figs. 10–12.

Firstly, Fig. 8 shows the significance and the effect of using the non-local total fractional-order variation regularizer and the diffusion regularizer. Carefully comparing Figs. 8(c)–(d), we notice that our new model deals with the MRI images better than the diffusion model. From the errors T_diff and T_frac between the registered image to the reference R in Figs. 9(a)–(c), we can see that our model reduces more efficiently the dissimilarity.

Secondly, the comparisons with some registration methods on a natural X-ray image are shown in Figs. 10–12. Note that both the template and the reference have some patches of noise. These noises cause serious trouble for the elastic and the linear curvature registration schemes. As is apparent from the picture sequence in Fig. 10(c)–(f), only the men curvature model produces visually correct registration results which look as good as from our model. However the respective displacement fields in Fig. 11 seem to suggest the opposite and, indeed, both Fig. 12 and its associated MFN values confirm that our fractional-order model is the best as it has no non-physical folding of meshes.

4.4. Fractional-power of the gradient versus fractional order derivative

Finally we show another comparison. In recent years, motivated by compressed sensing modeling, fractional-power of the gradient based regularizers has been used in many problems of scientific computing. Although this paper promotes the use of fractional order derivatives based regularizers, it is of interest to compare our model with the following fractional-power regularization image registration model
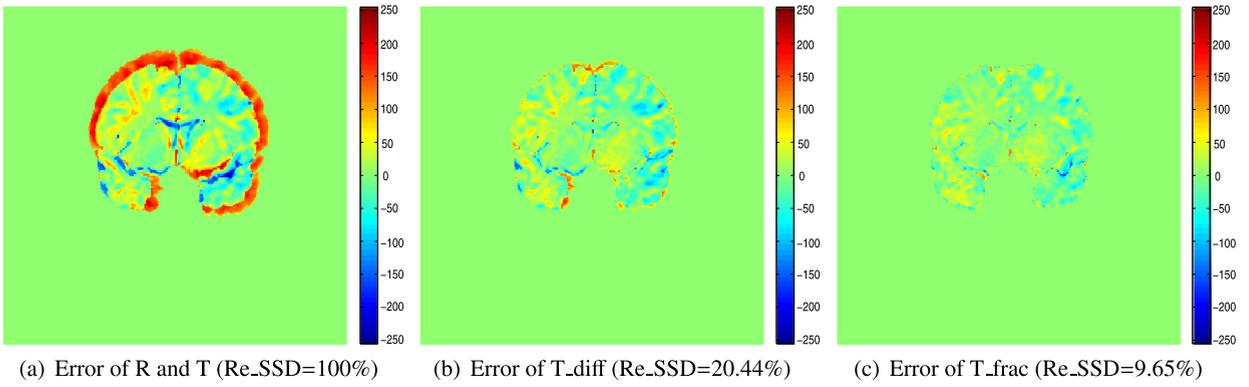
(a) Error of R and T (Re_SSD=100%)   (b) Error of T_diff (Re_SSD=20.44%)   (c) Error of T_frac (Re_SSD=9.65%)

Fig. 9. Dissimilarity presentations – comparisons between diffusion and our fractional.

(a) Reference R   (b) Template T   (c) Elastic (Re_SSD=1 43%)

(d) Linear curvature (Re_SSD=3.80%)   (e) Mean curvature (Re_SSD=0.77%)   (f) Our fractional (Re_SSD=6.28%)
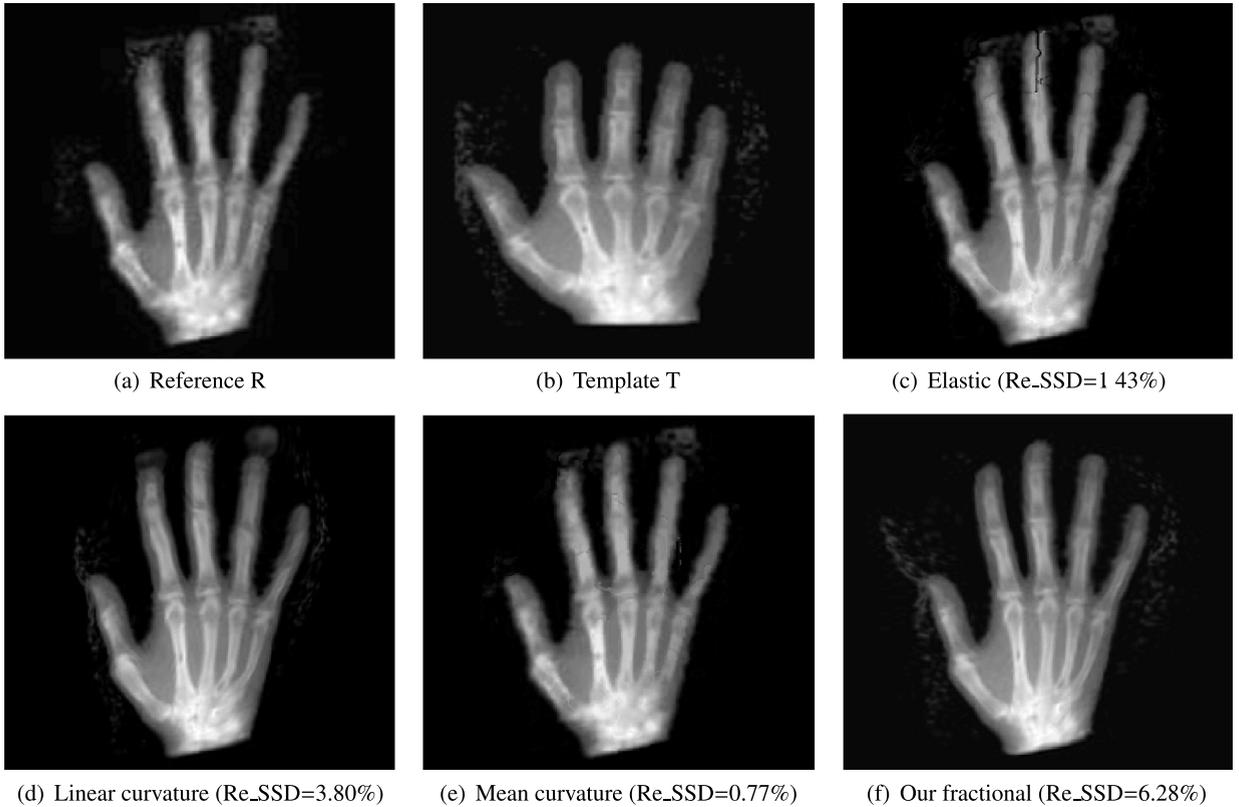
Fig. 10. Comparison of registered results of four different regularizers.

$$\min_{\boldsymbol{u}} \int_{\Omega} \sum_{i=1}^{2} |\nabla u_i|^{\beta} dx + \frac{\lambda}{2} \int_{\Omega} (T(x + \boldsymbol{u}(x)) - R(x))^2 dx, \tag{21}$$

using the same fraction β. For the Lena deformation data, Fig. 13 presents results showing difference images before and after registration. The difference images for fractional-power regularization registration ($\beta = 1.6$) are visibly poorer when our model as measured by the lowest registration error. Our proposed model with $\alpha = 1.6$ is able to recover the known deformations in this case. The test shows that, while fractional-power based regularizers produce reasonable results, fractional differentiation techniques could provide additional information when intensity contrast is insufficient or confounding [54].

5. Conclusions

The paper presented a total fractional-order variation based variational model for image registration. The use of a fractional order variation has improved unsteady deformation (e.g. not one to one) over competing models when the differences
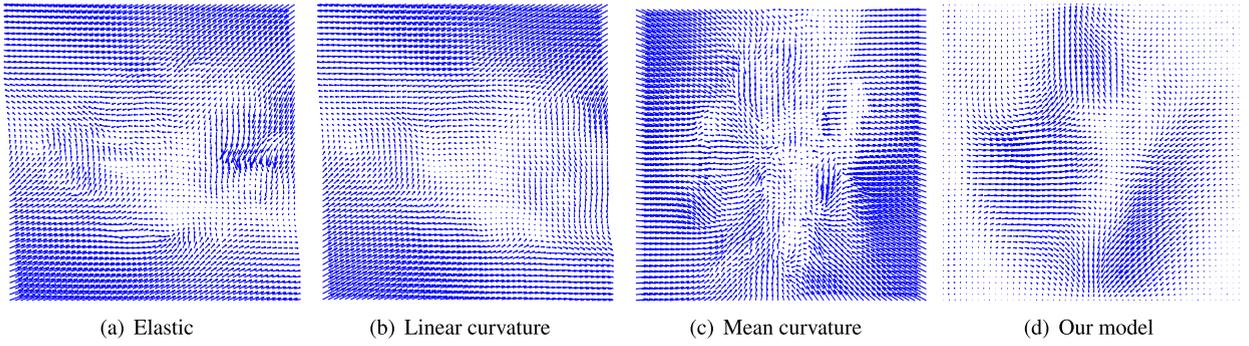
(a) Elastic   (b) Linear curvature   (c) Mean curvature   (d) Our model

Fig. 11. Comparison of displacement fields for different regularizers (flow map).

(a) Elastic
(MFN=3651,MFR=5.57%))

(b) Linear curvature
(MFN=56,MFR=0.09%))

(c) Mean curvature
(MFN=5210,MFR=7.95%))

(d) Our model
(MFN=0, MFR=0%)
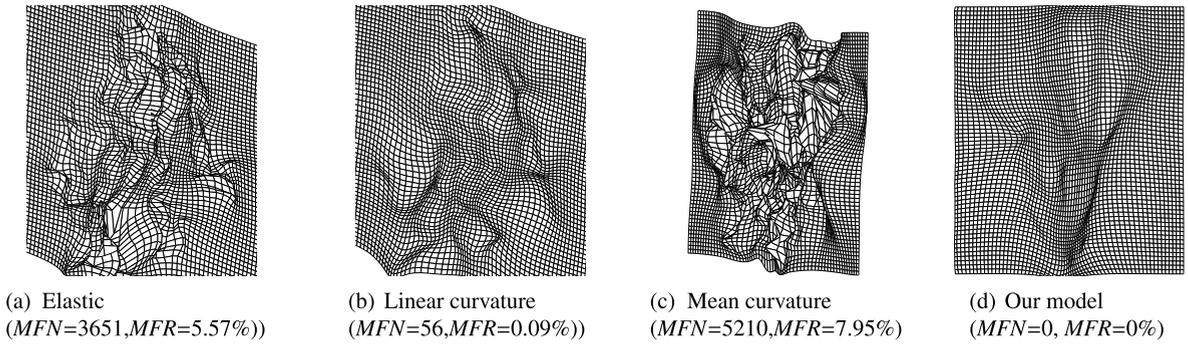
Fig. 12. Comparison of transformations $\phi(\boldsymbol{u}(x))$ for different regularizers (mesh plot).

(a) Reference   (b) Template   (c) Model (21) $\beta = 1.6$ (d) Our model $\alpha = 1.6$

(e) Error of R and T (Re_SSD=100%)   (f) Error of model (21) : Re_SSD=12.92%   (g) Error of our model: Re_SSD=4.40%
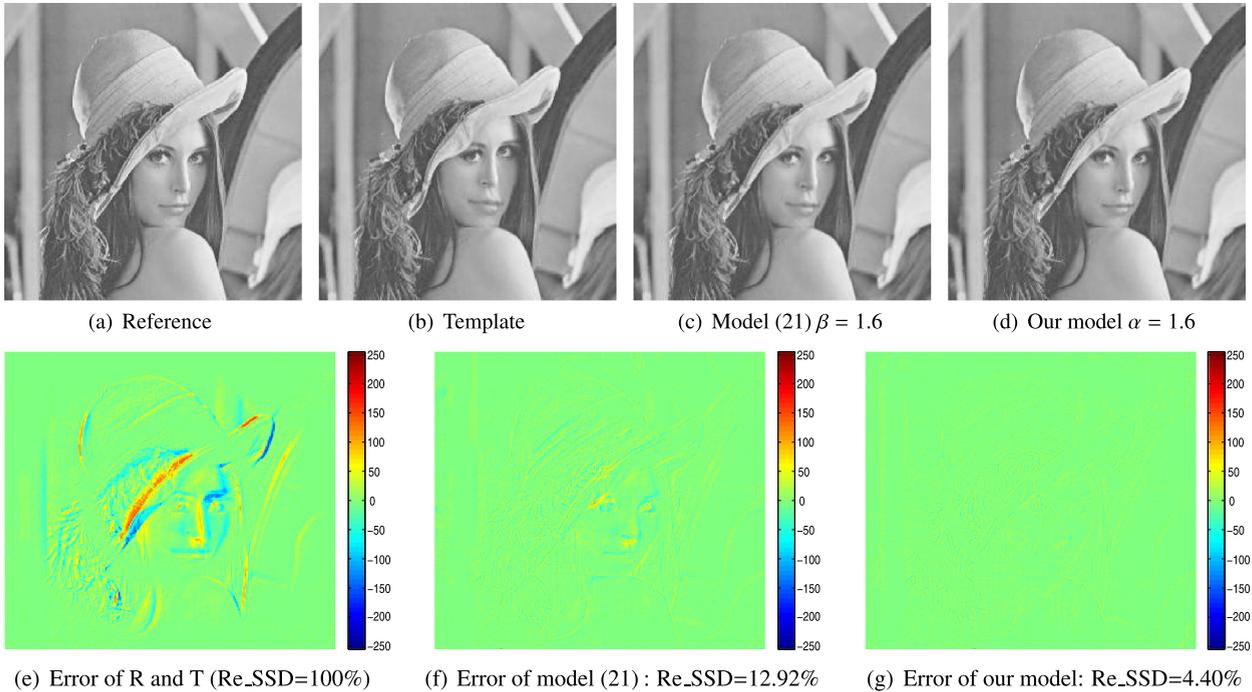
Fig. 13. Comparison between our model and model (21).

between the reference and template images are large. We focused on the theories, efficiency and application of the total fractional-order variation regularization with fractional order derivative and gave a numerical implementation of the new model combining the semi-implicit update and CG solution to the discretized matrix approximation systems of the underlying Euler–Lagrange equations. Numerical results show that the new model is more effective than the widely used elastic, diffusion and curvature regularizers based models. There is much scope that remains to be explored in applying fractional-order derivative to image applications.

References

[1] M.L. Kessler, Image registration and data fusion in radiation therapy, Br. J. Radiol. 79 (1) (2006) S99–S108, http://dx.doi.org/10.1259/bjr/70617164.
[2] S.K. Kyriacou, C. Davatzikos, S.J. Zinreich, R.N. Bryan, Nonlinear elastic registration of brain images with tumor pathology using a biomechanical model [mri], IEEE Trans. Med. Imaging 18 (7) (1999) 580–592.
[3] A. Mohamed, E.I. Zacharaki, D. Shen, C. Davatzikos, Deformable registration of brain tumor images via a statistical model of tumor-induced deformation, Med. Image Anal. 10 (5) (2006) 752–763.
[4] C. Nimsky, O. Ganslandt, P. Hastreiter, R. Fahlbusch, Intraoperative compensation for brain shift, Surg. Neurol. 56 (6) (2001) 357–364.
[5] J. Modersitzki, Numerical Methods for Image Registration, Numer. Math. Sci. Comput., Oxford University Press, USA, 2004.
[6] A.A. Goshtasby, 2-D and 3-D Image Registration: For Medical, Remote Sensing, and Industrial Applications, John Wiley & Sons, 2005.
[7] N. Chumchob, K. Chen, A robust affine image registration method, Int. J. Numer. Anal. Model. 6 (2) (2009) 311–334.
[8] N. Chumchob, K. Chen, C. Brito-Loeza, A fourth-order variational image registration model and its fast multigrid algorithm, Multiscale Model. Simul. 9 (1) (2011) 89–128.
[9] G.E. Christensen, M.I. Miller, M. Vannier, A 3d deformable magnetic resonance textbook based on elasticity, in: AAAI Spring Symposium Series: Applications of Computer Vision in Medical Image Processing, Citeseer, 1994, pp. 153–156.
[10] B. Fischer, J. Modersitzki, Fast diffusion registration, in: Contemp. Math., vol. 313, 2002, pp. 117–128.
[11] B. Fischer, J. Modersitzki, A unified approach to fast image registration and a new curvature based registration technique, Linear Algebra Appl. 380 (2004) 107–124.
[12] S. Henn, K. Witsch, A multigrid approach for minimizing a nonlinear functional for digital image matching, Computing 64 (4) (2000) 339–348.
[13] T. Vercauteren, X. Pennec, A. Perchant, N. Ayache, Diffeomorphic demons: efficient non-parametric image registration, NeuroImage 45 (1) (2009) S61–S72.
[14] A.S. Chaves, A fractional diffusion equation to describe Lévy flights, Phys. Lett. A 239 (1) (1998) 13–16.
[15] R. Metzler, J. Klafter, The random walk's guide to anomalous diffusion: a fractional dynamics approach, Phys. Rep. 339 (1) (2000) 1–77.
[16] I. Podlubny, Fractional Differential Equations: An Introduction to Fractional Derivatives, Fractional Differential Equations, to Methods of Their Solution and Some of Their Applications, Math. Sci. Eng., Elsevier Science, 1999.
[17] F.R. Lin, S.W. Yang, X.Q. Jin, Preconditioned iterative methods for fractional diffusion equation, J. Comput. Phys. 256 (2014) 109–117.
[18] H. Wang, N. Du, A superfast-preconditioned iterative method for steady-state space-fractional diffusion equations, J. Comput. Phys. 240 (2013) 49–57.
[19] H. Wang, N. Du, A fast finite difference method for three-dimensional time-dependent space-fractional diffusion equations and its efficient implementation, J. Comput. Phys. 253 (2013) 50–63.
[20] H. Wang, N. Du, Fast solution methods for space-fractional diffusion equations, J. Comput. Appl. Math. 255 (2014) 376–383.
[21] M. Zayernouri, G.E. Karniadakis, Fractional Sturm–Liouville eigen-problems: theory and numerical approximation, J. Comput. Phys. 252 (2013) 495–517.
[22] M. Zayernouri, G.E. Karniadakis, Exponentially accurate spectral and spectral element methods for fractional ODEs, J. Comput. Phys. 257 (2014) 460–480.
[23] J. Bai, X.C. Feng, Fractional-order anisotropic diffusion for image denoising, IEEE Trans. Image Process. 16 (10) (2007) 2492–2502, http://dx.doi.org/10.1109/tip.2007.904971.
[24] P. Guidotti, A new nonlocal nonlinear diffusion of image processing, J. Differ. Equ. 246 (12) (2009) 4731–4742.
[25] P. Guidotti, J.V. Lambers, Two new nonlinear nonlocal diffusions for noise reduction, J. Math. Imaging Vis. 33 (1) (2009) 25–37.
[26] M. Janev, S. Pilipović, T. Atanacković, R. Obradović, N. Ralević, Fully fractional anisotropic diffusion for image denoising, Math. Comput. Model. 54 (1) (2011) 729–741.
[27] P.D. Romero, V.F. Candela, Blind deconvolution models regularized by fractional powers of the Laplacian, J. Math. Imaging Vis. 32 (2) (2008) 181–191, http://dx.doi.org/10.1007/s10851-008-0093-2.
[28] D. Chen, S. Sun, C. Zhang, Y. Chen, D. Xue, Fractional-order TV-L^2 model for image denoising, Cent. Eur. J. Phys. (2013) 1–9, http://dx.doi.org/10.2478/s11534-013-0241-1.
[29] D. Chen, Y. Chen, D. Xue, Three fractional-order TV-L^2 models for image denoising, J. Comput. Inf. Syst. 9 (12) (2013) 4773–4780.
[30] J. Zhang, Z. Wei, L. Xiao, Adaptive fractional-order multi-scale method for image denoising, J. Math. Imaging Vis. 43 (1) (2012) 39–49, http://dx.doi.org/10.1007/s10851-011-0285-z.
[31] Y. Zhang, Y.F. Pu, J.R. Hu, J.L. Zhou, A class of fractional-order variational image inpainting models, Appl. Math. Inf. Sci. 6 (2) (2012) 299–306.
[32] A. Melbourne, N. Cahill, C. Tanner, M. Modat, D.J. Hawkes, S. Ourselin, Using fractional gradient information in non-rigid image registration: application to breast MRI, in: Medical Imaging 2012: Image Processing, Proc. SPIE 8314 (2012), http://dx.doi.org/10.1117/12.911173, 83141Z-83141Z-10.
[33] C.C. Garvey, N.D. Cahill, A. Melbourne, C. Tanner, S. Ourselin, D.J. Hawkes, Nonrigid image registration with two-sided space-fractional partial differential equations, in: ICIP, 2013, pp. 747–751.
[34] R. Verdú-Monedero, J. Larrey-Ruiz, J. Morales-Sánchez, J.L. Sancho-Gómez, Fractional regularization term for variational image registration, Math. Probl. Eng. 2009 (2009) 1–13.
[35] X. Zhang, Y. Shen, S. Li, H. Zhang, Medical image registration in fractional Fourier transform domain, Optik 124 (12) (2013) 1239–1242.
[36] J.P. Thirion, Image matching as a diffusion process: an analogy with Maxwell's demons, Med. Image Anal. 2 (3) (1998) 243–260.
[37] A. Godley, E. Ahunbay, C. Peng, X.A. Li, Automated registration of large deformations for adaptive radiation therapy of prostate cancer, Med. Phys. 36 (4) (2009) 1433–1441.
[38] H. Wang, L. Dong, J. O'Daniel, R. Mohan, A.S. Garden, K.K. Ang, D.A. Kuban, M. Bonnen, J.Y. Chang, R. Cheung, Validation of an accelerated 'demons' algorithm for deformable image registration in radiation therapy, Phys. Med. Biol. 50 (12) (2005) 2887.
[39] S. Henn, K. Witsch, Multimodal image registration using a variational approach, SIAM J. Sci. Comput. 25 (4) (2004) 1429–1447.
[40] K.B.A. Oldham, J.A. Spanier, The Fractional Calculus: Theory And Applications of Differentiation and Integration to Arbitrary Order, Dover Books Math., Dover Publications, Inc., 2006.
[41] O.P. Agrawal, Fractional variational calculus in terms of Riesz fractional derivatives, J. Phys. A, Math. Theor. 40 (24) (2007) 62–87.
[42] R. Almeida, D.F.M. Torres, Necessary and sufficient conditions for the fractional calculus of variations with Caputo derivatives, Commun. Nonlinear Sci. Numer. Simul. 16 (3) (2011) 1490–1500, http://dx.doi.org/10.1016/j.cnsns.2010.07.016.

[43] T.M. Atanackovic, S. Konjik, S. Pilipovic, Variational problems with fractional derivatives: Euler–Lagrange equations, J. Phys. A, Math. Theor. 41 (9) (2008) 095201.
[44] A.A.A. Kilbas, H.M. Srivastava, J.J. Trujillo, Theory and Applications of Fractional Differential Equations, North-Holl. Math. Stud., Elsevier Science & Tech, 2006.
[45] L. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms, Phys. D, Nonlinear Phenom. 60 (1) (1992) 259–268.
[46] R. Hilfer, Applications of Fractional Calculus in Physics, vol. 5, World Scientific, 2000.
[47] J.P. Zhang, K. Chen, A total fractional-order variation model for image restoration with non-homogeneous boundary conditionals and its numerical solution, submitted for publication.
[48] L.C. Evans, R.F. Gariepy, Measure Theory and Fine Properties of Functions, Stud. Adv. Math., CRC Press, Boca Raton, FL, 1992.
[49] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, F. Lenzen, Variational Methods in Imaging, Appl. Math. Sci., Springer, 2009.
[50] R. Acar, C.R. Vogel, Analysis of bounded variation penalty methods for ill-posed problems, Inverse Probl. 10 (6) (1994) 1217–1229.
[51] G. Aubert, P. Kornprobst, Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations, 2nd edition, Appl. Math. Sci., vol. 147, Springer-Verlag, 2006.
[52] M.M. Meerschaert, C. Tadjeran, Finite difference approximations for two-sided space-fractional partial differential equations, Appl. Numer. Math. 56 (1) (2006) 80–90, http://dx.doi.org/10.1016/j.apnum.2005.02.008.
[53] J.P. Zhang, K. Chen, B. Yu, An iterative Lagrange multiplier method for constrained total-variation-based image denoising, SIAM J. Numer. Anal. 50 (3) (2012) 983–1003, http://dx.doi.org/10.1137/110829209.
[54] B. Mathieu, P. Melchior, A. Oustaloup, C. Ceyral, Fractional differentiation for edge detection, Signal Process. 83 (11) (2003) 2421–2432.