Noname manuscript No.
(will be inserted by the editor)

Nonlocal elastica model for sparse reconstruction

Mengyuan Yan · Yuping Duan

Received: date / Accepted: date

Abstract In view of the exceptional ability of curvature in connecting missing edges and structures, we propose novel sparse reconstruction models via the Euler's elastica energy. Especially, we firstly extend the Euler's elastica regularity into the nonlocal formulation to fully take the advantages of the pattern redundancy and structural similarity in image data. Due to its non-convexity, non-smoothness and nonlinearity, we regard both local and nonlocal elastica functional as the weighted total variation for a good trade-off between the runtime complexity and performance. The splitting techniques and alternating direction method of multipliers (ADMM) are used to achieve efficient algorithms, the convergence of which is also discussed under certain assumptions. The weighting function occurred in our model can be well estimated according to the local approach. Numerical experiments demonstrate that our nonlocal elastica model achieves the state-of-the-art reconstruction results for different sampling patterns and sampling ratios, especially when the sampling rate is extremely low.

Keywords Euler's elastica · nonlocal regularization · sparse reconstruction · ADMM

1 Introduction

The Euler's elastica was first introduced to computer vision by Mumford [36] as a prior curve model, where "elastica" means the shape of the ideal thin elastic rod on a plane. Observing the importance of level lines for image representation, Masnou *et al.* [34,1] proposed the variational formulation of the Euler's elastica by measuring the total length and total curvature of a planar curve $\mathcal{C}$ as follows

$$E(\mathcal{C}) = \int_{\mathcal{C}} (a + b\kappa^2(s))ds, \tag{1}$$

where $a, b > 0$ are two parameters, and κ is the curvature of $\mathcal{C}$ at position s. Thus, the Euler's elastica of an image $u : \Omega \to \mathbb{R}$, $\Omega \subset \mathbb{R}^2$ being an open bounded subset of $\mathbb{R}^2$ with Lipschitz continuous boundary, becomes

$$E(u) = \int_{\Omega} \left(a + b\kappa^2\right) |\nabla u| dx, \tag{2}$$

Mengyuan Yan
The Center for Applied Mathematics, Tianjin University, Tianjin, China 300072.

Y. Duan
The Center for Applied Mathematics, Tianjin University, Tianjin, China 300072.
E-mail: yuping.duan@tju.edu.cn

where κ is the curvature of the level lines of u defined as

$$\kappa(u) := \nabla \cdot \frac{\nabla u}{|\nabla u|}.$$

Owing to the strong priors for the continuity of edges, the Euler's elastica (2) is originally introduced as the regularization for image inpainting [11,50], which is used to interpolate the true image from the observed data with a missing or damaged portion. The Euler's elastica regularized variational model for image inpainting is then formulated as follows

$$\min_u \ \frac{\lambda}{2}\int_{\Omega\setminus D}(u-u_0)^2dx + \int_\Omega (a+b\kappa^2)|\nabla u|dx, \tag{3}$$

where $D \subset \Omega$ denotes the missing region, u_0 is the available part in $\Omega\setminus D$ and λ is a positive parameter to balance the regularization term and the fidelity term. It is well-known that the reconstructed image from the total variation (TV) based models [34] suffers from the staircasing effect such that piecewise constant solutions are favored instead of piecewise smooth ones. One main advantage of the Euler's elastica energy is that it gives more natural appearance by minimizing the total curvature as well as lengths of the level lines, which can overcome the staircasing effect. Due to the great success in image inpainting, the Euler's elastica has been applied to other image processing tasks, such as segmentation [54,16,5,26], segmentation with depth [18,53], illusory contour [35,28] and denoising [43,15], etc.

Indeed, the inpainting problem also exists in medical imaging. The well-known compressed sensing MRI (CS-MRI) deals with data interpolation in the k-space [42]. The CS-MRI methods enable to reconstruct high quality MR images based on the partial Fourier measurements, which can reduce the scanning time and avoid the motion artifacts happening in traditional scanning. Lustig, Donoho and Pauly [32,33] exploited the sparsity of MR images in certain transform domains, i.e., the spatial finite differences and their wavelet coefficients, and employed the L^1 minimization to realize such sparsity

$$\min_u \ \frac{\lambda}{2}\int_\Omega (Ku-g)^2dx + \alpha\int_\Omega |\mathcal{W}u|dx + \beta\int_\Omega |\nabla u|dx, \tag{4}$$

where the operator $K \in \mathbb{C}^{M\times N}$ with the sample size M being extremely small compared to the image size N (i.e., $M \ll N$), $\mathcal{W}$ is an orthogonal wavelet transform, g is the observed k-space data and α, β are two positive parameters.

In order to improve the reconstruction quality, different regularization techniques have been explored for CS-MRI. Guo and Yin [23] proposed an edge guided compressed sensing reconstruction approach by exploiting weighted total variation (TV) and wavelet transform as the regularization, where the edge detectors were used to improve the edge recovery. Zhang *et al.* [51] used the nonlocal (NL) TV as the regularization term for sparse reconstruction. Huang and Yang [27] proposed a variational model based on the wavelet sparsity and periodic NLTV regularization. Qu *et al.* [39] proposed a patch-based directional wavelets to reconstruct images from undersampled k-space data, which used a parameter to indicate the geometric direction of each patch. Later the authors designed the patch-based nonlocal operator to model the sparse representation of similar image patches in [40]. Compton *et al.* [14] developed a hybrid regularization model based on TV and framelets for MRI reconstruction when the k-space data is undersampled. Considering that real-life images are not piecewise constant and detail information may be lost using TV based methods. High order regularization techniques have also been explored for CS-MRI. Knoll *et al.* [29] used the second-order total generalized variation (TGV) [7] as the regularity for piecewise smooth MRI reconstruction. Guo *et al.* [22] proposed a new regularization scheme by integrating second-order TGV and shearlet transform for better recovery of the edges and fine details in images. Besides, other efficient denoising techniques and priors as well as deep convolution neural network (CNN) have been also introduced for CS-MRI. For instance, Ravishankar and Bresler [41] presented a framework for simultaneously learning the dictionary and reconstructing the image from highly undersampled k-space data. Panić *et al.* [38] proposed a CS-MRI reconstruction algorithm by making use of a Markov

random field prior model for spatial clustering of subband data. Eksioglu [17] developed a reconstruction algorithm using a decoupled iterations alternating over a denoising step realized by blocking matching 3D (BM3D) image model and a reconstruction step through an optimization formulation. Hammernik *et al.* [25] introduced an efficient trainable formulation for CS-MRI, which implemented a variational network to learn a complete reconstruction procedure. Yan *et al.* [48] reformulated an ADMM algorithm for CS-MRI into a deep CNN by learning the parameters end-to-end in the training phase. Yang *et al.* [47] proposed a conditional generative adversarial networks-based model for CS-MRI.

In this paper, we make two main contributions for sparse reconstruction. The **first contribution** is to investigate the application of the Euler's elastica energy for CS-MRI. We show the power of Euler's elastica regularization in connecting the missing level sets into smooth curves. More importantly, our **second contribution** is to define the curvature with nonlocal operators and minimize the following nonlocal elastica model to exploit the similarity between patches in the same image

$$\min_u \ \frac{\lambda}{2}\int_\Omega (Ku-g)^2 dx + \int_\Omega (a+b\kappa_{\mathrm{NL}}^2)|\nabla_{\mathrm{NL}} u| dx, \tag{5}$$

where both curvature and TV are defined by nonlocal operators. The proposed nonlocal elastica model can not only preserve the texture structure but also keep the spatial dependency of images. Along with the preprocessing data to compute the weight function, it is proved that the nonlocal extension of Euler's elastica provides an effective mechanism for the inverse problems. The numerical experiments on CS-MRI reconstruction demonstrate that the proposed nonlocal elastica regularization is superior to both the Euler's elastica and NLTV, regardless of low or high sampling rates.

The outline of the rest of the paper is as follows. In Sect. 2, we propose the Euler's elastica model for sparse reconstruction and discuss the numerical algorithms for minimizing the associated Euler's elastica functionals. Section 3 reviews the nonlocal methods and then extend the Euler's elastica model into the nonlocal formulation. An efficient ADMM based algorithm is also presented with a theoretical convergence analysis. Numerical experiments on synthetic and real image data are conducted in Sect. 4, where we compare the results against the state-of-the-art methods from the literature. We conclude the paper with remarks in Sect. 5.

2 Sparse reconstruction via Euler's elastica model

Motivated by the success of the Euler's elastica for image inpainting problem, we firstly utilize it as the regularization for the CS-MRI, which can be regarded as a k-space inpainting task. That is, we propose the following high-order variational model

$$\min_u \ \frac{\lambda}{2}\int_\Omega (Ku-g)^2 dx + \int_\Omega (a+b\kappa^2)|\nabla u| dx. \tag{6}$$

The minimization of (6) is numerically challenging due to the high non-smoothness, non-convexity and nonlinearity of the curvature term in the functional. Focused on the sparse reconstruction problem (6), we review three main numerical algorithms for Euler's elastica model as follows.

The CKS method The first numerical attempt to directly solve the Euler's elastica model was given by Chan, Kang and Shen in [11] based on the calculus of variation and the steepest descent which is derived for (6) as follows

$$\partial_t u(x,t) = \nabla\cdot V + \lambda K^*(Ku-g). \tag{7}$$

In (7), the vector field V is called the flux of the elastica energy,

$$V = (a+b\kappa^2)\frac{\nabla u}{|\nabla u|} - \frac{2b}{|\nabla u|^3}\nabla^\perp u \nabla(\kappa|\nabla u|)\nabla^\perp u.$$

Due to the restricted CFL condition, the convergence of the above gradient scheme requires a very small time step size, which can be accelerated using the optimal gradient method developed by Nesterov [37].

The THC method Thanks to the operator splitting and alternating direction methods, Tai, Hahn and Chung proposed a fast algorithm for Euler's elastica minimization problems in [43]. More specifically, by introducing the new variables p, n and m, the minimization (6) is rewritten as the following constrained optimization problem

$$\begin{aligned} \min_{u,p,m,n} \ & \frac{\lambda}{2}\int_\Omega (Ku-g)^2 dx + \int_\Omega \big(a+b(\nabla\cdot n)^2\big)|p|dx \\ \text{s.t.}, \ & p=\nabla u,\ |p| = m\cdot p,\ n=m,\ |m|\le 1. \end{aligned} \tag{8}$$

Based on the augmented Lagrangian method, the constrained optimization problem can be reformulated as the Lagrangian functional

$$\begin{aligned} \mathcal{L}_{\text{THC}}(u,p,m,n;\Lambda_1,\Lambda_2,\Lambda_3) = & \frac{\lambda}{2}\int_\Omega (Ku-g)^2 dx + \int_\Omega \big(a+b(\nabla\cdot n)^2\big)|p|dx + \int_\Omega \Lambda_1(p-\nabla u)dx \\ & + \frac{r_1}{2}\int_\Omega (p-\nabla u)^2 dx + \int_\Omega \Lambda_2(|p|-m\cdot p)dx + \frac{r_2}{2}\int_\Omega (|p|-m\cdot p)dx \\ & + \int_\Omega \Lambda_3(m-n)dx + \frac{r_3}{2}\int_\Omega (m-n)^2 dx + \delta_{\mathcal{R}}(m), \end{aligned}$$

where $\Lambda_1,\Lambda_2,\Lambda_3$ are Lagrange multipliers, r_1,r_2,r_3 are positive penalty parameters, and $\delta_{\mathcal{R}}(\cdot)$ denotes a characteristic function on $\mathcal{R}$. Although all variables can be solved by either fast Fourier Transform (FFT) or closed-form solution, it is not easy to choose the three model parameters and four penalty parameters and the convergence rate is also slow in practice.

The κTV method Letting $\phi(\kappa)=\big(a+b\kappa^2\big)$, Bae, Shi and Tai [4] treated $\phi(\kappa)$ as the weight for total variation and updated $\phi(\kappa)$ iteratively, which makes graph cuts applicable for solving the Euler's elastica denoising model. Yashtini and Kang [50] rewrote the Euler's elastica model for image inpainting as a weighted TV model in exactly the same way and called it the κTV model, such that many fast solvers can be used for the numerical implementation such as Chambolle's dual method [9], primal-dual method [10], splitting Bregman method [21], and augmented Lagrangian method [46].

More specifically, by introducing an auxiliary variable p, the following constrained minimization problem is considered with κ being assumed known in advance

$$\begin{aligned} \min_{u,p} \ & \frac{\lambda}{2}\int_\Omega (Ku-g)^2 dx + \int_\Omega \phi(\kappa)|p|dx \\ \text{s.t.}, \ & p=\nabla u. \end{aligned} \tag{9}$$

By applying the augmented Lagrangian method, we achieve the following saddle-point problem

$$\min_{u,p}\max_{\Lambda} \ \mathcal{L}_{\kappa\text{TV}}(u,p;\Lambda) := \frac{\lambda}{2}\int_\Omega (Ku-g)^2 dx + \int_\Omega \phi(\kappa)|p|dx + \int_\Omega \Lambda(p-\nabla u)dx + \frac{r}{2}\int_\Omega (p-\nabla u)^2 dx,$$

where Λ is a Lagrange multiplier and r is a positive parameter. Although the ground truth of $\phi(\kappa)$ is unknown for general inverse problems such as sparse reconstruction, it can be approximated using either the observed data or the most recent solution of u generated by the iterative scheme similar to [4,50]. We sketch the proximal ADMM for the above saddle-point problem as Algorithm 1, where a proximal term is introduced in the u-subproblem to guarantee the convergence of algorithm in theory.

There is the following theoretical guarantee for the convergence of the κTV method. We omit the detailed proof since it is almost the same as Theorem 3 in [50].

Algorithm 1: The κTV method

1: **Initialize:** $u^0=0$, $p^0=0$, $\Lambda^0=0$, $k=0$, ϵ, ς, $nIter$;
2: **While** $k<nIter$ and $\|u^{k+1}-u^k\|_2^2<\epsilon\|u^k\|_2^2$
3: Solve u^{k+1} from

$$u^{k+1}=\arg\min_u \frac{\lambda}{2}\int_\Omega (Ku-g)^2dx+\frac{r}{2}\int_\Omega\Big(\nabla u-p^k-\frac{\Lambda^k}{r}\Big)^2dx+\frac{\gamma}{2}\int_\Omega(u-u^k)^2dx, \tag{10}$$

which is solved by fast Fourier transform (FFT).
4: Update $\phi(\kappa^{k+1})$ from

$$\phi(\kappa^{k+1})=a+b\left(\nabla\cdot\frac{\nabla u^{k+1}}{|\nabla u^{k+1}|_\varsigma}\right)^2 \text{ with } |\nabla u^{k+1}|_\varsigma=\max\{\varsigma,|\nabla u^{k+1}|\}, \tag{11}$$

here, $\varsigma>0$ is used to avoid the denominator vanishing.
5: Solve p^{k+1} from

$$p^{k+1}=\arg\min_p \int_\Omega\phi(\kappa^{k+1})|p|dx+\frac{r}{2}\int_\Omega\Big(p-\nabla u^{k+1}+\frac{\Lambda^k}{r}\Big)^2dx, \tag{12}$$

which is solved by shrinkage operator.
6: Update Λ^{k+1} from

$$\Lambda^{k+1}=\Lambda^k+r(p^{k+1}-\nabla u^{k+1}). \tag{13}$$

7: **end while**

Theorem 1 *Let $\{(u^k,p^k;\Lambda^k)\}_{k\in\mathbb{N}}$ be the sequence generated by the Algorithm 1 and $(\bar u,\bar p;\bar\Lambda)$ be a point satisfying the first-order optimality conditions.*

$$\begin{cases}\lambda K^*(Ku-g)+\mathrm{div}\Lambda=0,\\ \phi(\kappa)\partial|p|+\Lambda\ni 0,\quad \text{where } \phi(\kappa)=a+b\big(\nabla\cdot\frac{\nabla u}{|\nabla u|}\big)^2,\\ p-\nabla u=0.\end{cases} \tag{14}$$

For any $\tau^k\in\partial|p^k|$, we define

$$\Theta^k=\Big\langle\big(\phi(\kappa^k)-\phi(\bar\kappa)\big)\tau^k,p^k-\bar p\Big\rangle,\quad \forall k\in\mathbb{N}. \tag{15}$$

Assume there exists a solution to problem (9) *and Θ^k is always nonnegative for any k. Then the sequence $\{(u^k,p^k;\Lambda^k)\}_{k\in\mathbb{N}}$ converges to a limit point that satisfies the first-order optimality conditions* (14).

Proof The proof follows the idea of Theorem 3 in [50]. It should be mentioned that there are two differences between the one in [50] and ours. First, we introduce one artificial variable while Theorem 3 in [50] used two artificial variables for solving the inpainting problem. Second, the operator K represents a subsampled Fourier transform while it is a selection operator for the inpainting model. Indeed, it does not affect the proof of convergence because K is semi-definite in both cases. □

Remark 1 We simply let $\gamma=0$ in the numerical experiments and observe that the Θ^k is always nonnegative numerically as shown in Fig. 7.

3 Nonlocal elastica model

3.1 Nonlocal approaches

It is well-known that TV minimization favors piecewise constant solutions, which regards the image details such as textures and repetitive structures as noises in denoising processing. Motivated by the

pioneer work [49,44], Buades, Coll and Morel [8] proposed the nonlocal means filter, which introduced the patch distance to use the value at point y for denoising x if the weighted average of pixels of the patch centered at y is similar to the local patch centered at x. Later, Gilboa and Osher [20] embedded the nonlocal means into the variational formulation, which facilitates the applications of the nonlocal regularity for general inverse problems such as image segmentation [19] and reconstruction [51].

Nonlocal operators In order to define the nonlocal elastica, we follow the variational formulation of nonlocal operators in [20]. Let $w(x,y):\Omega\times\Omega\to\mathbb{R}^+$ be a weight function for $x,y\in\Omega$, which is nonnegative and symmetric. The nonlocal gradient $\nabla_w u:\Omega\to\Omega\times\Omega$ is defined as the vector for all pairs of pixels $x,y\in\Omega$

$$\nabla_w u(x,y)=(u(y)-u(x))\sqrt{w(x,y)}.$$

Let p_1, p_2 be the nonlocal vectors such that $p_i:\Omega\times\Omega\to\mathbb{R}$ for $i=1,2$. The inner product between p_1 and p_2 is defined as

$$\langle p_1,p_2\rangle=\int_\Omega p_1(x,y)p_2(x,y)dy.$$

Accordingly, we can define the norm of a nonlocal vector $p:\Omega\times\Omega\to\mathbb{R}$ at point $x\in\Omega$ as

$$|p|(x)=\sqrt{\int_\Omega p(x,y)^2dy}.$$

Due to the following joint relation between the nonlocal gradient and divergence for $u:\Omega\to\mathbb{R}$ and $p:\Omega\times\Omega\to\mathbb{R}$

$$\langle\nabla_w u,p\rangle=-\langle u,\mathrm{div}_w p\rangle,$$

the nonlocal divergence operator can be given as

$$\mathrm{div}_w p(x)=\int_\Omega (p(x,y)-p(y,x))\sqrt{w(x,y)}dy.$$

Furthermore, the graph Laplacian is defined using the gradient and divergence operator as follows

$$\Delta_w u(x)=\frac{1}{2}\mathrm{div}_w(\nabla_w u(x))=\int_\Omega (u(y)-u(x))w(x,y)dy,$$

where the factor $\frac{1}{2}$ is used to get the standard Laplacian definition. It is easy to check that the nonlocal Laplacian operator is self-adjoint and negative semi-definite as the local Laplacian operator.

Another important issue of the nonlocal methods is how to define the weight function between pixels, which is used to measure the similarity of image features between any two pixels $x\in\Omega$ and $y\in\Omega$. Similarly to [8], we estimate the weight function as follows

$$w(x,y)=\exp\left(-\frac{\sum\limits_{z\in q(o)}G_a(z)\big|f(x+z)-f(y+z)\big|^2}{2h^2}\right),\tag{16}$$

where G_a is the Gaussian kernel with standard deviation a, z is the pixel in $q(o)$, a squared patch centered in origin and h is the filtering parameter.

NonLocal TV Gilboa and Osher [20] proposed the image denoising model using the nonlocal TV norm, which recovers the noisy image based on the L^1 norm of nonlocal gradient

$$\min_u \ \frac{\lambda}{2}\int_\Omega (u-f)^2 dx + \int_\Omega |\nabla_w u| dx. \tag{17}$$

Similarly to the classical TV model, i.e., the ROF model, different numerical algorithms have been applied to solve the NLTV minimization problem. For instance, Gilboa and Osher extended the Chambolle's dual algorithm for the NLTV, which is much faster than the direct steepest descent scheme. Zhang *et al.* [51] developed the split Bregman algorithm for the NLTV based sparse reconstruction model.

Although the nonlocal methods show superior performance in texture and structure repeated image processing tasks, the restoration results lack spatial dependency due to the point-to-point estimation. Indeed, natural images are usually high structured as their pixels exhibit strong spatial dependency. In Fig. 1, we apply the Euler's elastica model, NLTV and nonlocal Euler's elastica (NLEE) model to the 'Cameraman' image corrupted by additive Gaussian noise of mean 0 and standard variance 0.1. The mean squared error of the NLTV is smaller (i.e., high PSNR) than the Euler's elastica, but it lacks spatial smoothness. Higher SSIM indicates that the high-order regularization can preserve better spatial correlation. Obviously, the nonlocal elastica regularization is shown to be able to remove noises as well as preserve the structure similarity. Thus, we introduce the nonlocal curvature and minimize the nonlocal elastica energy for sparse reconstruction problems in forthcoming subsection.

(a) 20.00dB/0.3286 (b) 26.87dB/0.7904 (c) 27.34dB/0.7311 (d) 27.77dB/0.8024

Fig. 1 The test for denoising. From left to right: noise image, THC, NLTV and NLEE, where $\cdot/\cdot$ denotes the value of PSNR and SSIM, respectively

3.2 Nonlocal elastica model for sparse reconstruction

Computing the weights The weight function for image denoising can be precomputed and fixed since most similarity information is available in the noisy image. It is difficult to obtain a good estimation of the weight function for image inpainting and reconstruction because a large quantity of information is missing and the given data even lie in the transformed domain. So far, various methods have been proposed to estimate the weight function. Arias *et al.* [2] regarded the weight function as unknown and re-estimated it from the current image estimate. Later, the authors [3] proposed a block based cost functional by introducing an entropy regularization to adaptively estimate the weight function. Yang and Jacob [52] used the majorize-minimize framework to develop a two-step alternating scheme, which is to perform the classical H^1-NL algorithm followed by the re-estimation of the weights from the current image estimation.

It is straightforward to use similar strategies to update the weight function according to the latest estimation for our nonlocal elastica model. However, the nonlocal approaches usually converge slowly with

the updated weights (e.g., [51]), which greatly increases the computational burden. Both the imaging problem and high-order regularization in our case require fast numerical solvers. Thus, we use a two-stage method by using the reconstructed data of the κTV model (6) as an initial image guess to compute the weights for nonlocal elastica model up front. Similar technique has been used in [31], which used the Filtered Back Projection (FBP) image to compute the weights for the NLTV-based tomographic reconstruction. In what follows, we use fixed weights in the nonlocal elastica model for sparse reconstruction problem, which can not only improve the computation and storage efficiency but also provide reconstruction results with high quality. The numerical comparison on the nonlocal elastica model with updated and fixed weight function is conducted in the Sect. 4.3, which demonstrates the two-stage framework works well for different sampling patterns.

Nonlocal elastica Per the previous discussion, we reformulate the Euler's elastica model (6) into the following nonlocal version

$$\min_u \ \frac{\lambda}{2}\int_\Omega (Ku-g)^2dx + \int_\Omega \Big(a + b\Big(\mathrm{div}_w\big(\frac{\nabla_w u}{|\nabla_w u|}\big)\Big)^2\Big)|\nabla_w u|dx. \tag{18}$$

Obviously, it requires to solve a fourth-order Partial Differential Equation (PDE) by using the gradient descent, which is extremely time consuming for such a nonlocal model. Inspired by the κTV method, we regard the nonlocal elastica model (18) as a re-weighted NLTV model, called the NLκTV model, such that

$$\min_u \ \frac{\lambda}{2}\int_\Omega (Ku-g)^2dx + \int_\Omega \phi(\kappa_w)|\nabla_w u|dx, \tag{19}$$

where $\phi(\kappa_w) = a + b\kappa_w^2$ with κ_w being the nonlocal curvature, i.e., $\kappa_w = \mathrm{div}_w\big(\frac{\nabla_w u}{|\nabla_w u|}\big)$. Similarly, we use the ADMM to solve the minimization problem by introducing an auxiliary variable v, and rewrite (19) into the following constrained minimization problem

$$\begin{aligned} &\min_{u,v} \ \frac{\lambda}{2}\int_\Omega (Kv-g)^2dx + \int_\Omega \phi(\kappa_w)|\nabla_w u|dx \\ &\text{s.t.,} \ \ u = v. \end{aligned} \tag{20}$$

Analogously, we reformulate the constrained minimization as a saddle-point problem using the augmented Lagrangian method such that

$$\min_{u,v}\max_{\Lambda_1} \mathcal{L}_{\mathrm{NL}\kappa\mathrm{TV}}(u,v;\Lambda_1) := \frac{\lambda}{2}\int_\Omega (Kv-g)^2dx + \int_\Omega \phi(\kappa_w)|\nabla_w u|dx + \int_\Omega \Lambda_1(u-v)dx + \frac{r_1}{2}\int_\Omega (u-v)^2dx,$$

with Λ_1 being an associated Lagrange multiplier, and r_1 being a positive parameter. The alternating minimization algorithm is described as Algorithm 2.

We are now left with the minimization problem of the variable v and u corresponding to the reconstruction problem and the nonlocal denoising problem, respectively. We detail the solutions as follows.

The sub-minimization w.r.t. v For v-subproblem, we consider the following minimization problem

$$\min_v \ \frac{\lambda}{2}\int_\Omega (Kv-g)^2dx - \int_\Omega \Lambda_1^k v dx + \frac{r_1}{2}\int_\Omega (v-u^k)^2dx + \frac{\gamma}{2}\int_\Omega (v-v^k)^2dx,$$

where the proximal term is introduced to guarantee the convergence. The first-order optimality condition of the above minimization gives us a linear problem, i.e.,

$$\big(\lambda K^*K + (r_1+\gamma)\mathcal{I}\big)v = \lambda K^*g + r_1u^k + \gamma v^k + \Lambda_1^k, \tag{21}$$

Algorithm 2: The two-stage NLκTV method

Input: $g, \epsilon, \varsigma, nIter$.
Output: u.
/* First-stage */
κTV reconstruction: Implement Algorithm 1 to estimate the reconstructed image denoted as $u^{\kappa TV}$;
/* Second-stage */
$u^0 \leftarrow u^{\kappa\mathrm{TV}}$; $v^0 \leftarrow u^0$;
Compute the weight function: Calculate the nonlocal weight function according to (16) based on the $u^{\kappa TV}$;
for $k = 0$ *to* $nIter - 1$ **do**
  $v^{k+1} \leftarrow \arg\min_v \mathcal{L}_{\mathrm{NL}\kappa\mathrm{TV}}(u^k, v; \Lambda_1^k)$; /* solve the reconstruction problem */
  $u^{k+1} \leftarrow \arg\min_u \mathcal{L}_{\mathrm{NL}\kappa\mathrm{TV}}(u, v^{k+1}; \Lambda_1^k)$; /* solve the nonlocal denosing problem */
  $\phi(\kappa_w^{k+1}) \leftarrow a + b\left(\mathrm{div}_w\left(\frac{\nabla_w u^{k+1}}{|\nabla_w u^{k+1}|_\varsigma}\right)\right)^2$ with $|\nabla_w u^{k+1}|_\varsigma = \max\{\varsigma, |\nabla_w u^{k+1}|\}$;
  /* let $\varsigma > 0$ to avoid the denominator vanishing */
  $\Lambda_1^{k+1} \leftarrow \Lambda_1^k + r_1(u^{k+1} - v^{k+1})$;
  if $\|u^{k+1} - u^k\|_2^2 < \epsilon\|u^k\|_2^2$ **then**
    break;
  end
end

where $\mathcal{I}$ denotes the identity operator and K^* denotes the conjugate transpose operator of K. For CS-MRI, K represents the subsampled Fourier transform, which is defined as $K = PF$ with $P \in \mathbb{R}^{M\times N}$ being a selection matrix and F being a 2D matrix representing 2D discrete Fourier transform. The vector K^*g is zero-filling-based reconstructed image for the observed Fourier data g. Because $F^*F = \mathcal{I}$, the left-hand side operator of (21) can be diagonalized by the discrete Fourier transform F such as

$$F\big(\lambda K^*K + (r_1+\gamma)\mathcal{I}\big)F^* = \lambda P^TP + (r_1+\gamma)\mathcal{I}.$$

Thus, there is the following closed-form solution to the variable v based on the discrete Fourier transform

$$v^{k+1} = F^*\Big(\big(\lambda P^Tg + F\big(r_1u^k + \gamma v^k + \Lambda_1^k\big)\big)/\big(\lambda P^TP + (r_1+\gamma)\mathcal{I}\big)\Big), \tag{22}$$

the denominator of which can be precomputed at the beginning of the algorithm and used repeatedly during the iteration process.

The sub-minimization w.r.t. u The concrete form of the u-sub problem is given as follows

$$\min_u \int_\Omega \phi(\kappa_w^k)|\nabla_w u|dx + \int_\Omega \Lambda_1^k u dx + \frac{r_1}{2}\int_\Omega (u - v^{k+1})^2 dx, \tag{23}$$

which is somehow difficult to solve due to the non-differentiability of the TV semi-norm. Thus, we use the operator splitting technique to decouple the non-differentiable term and the squared L^2-norm term. By introducing another variable p, we can rewrite the minimization problem into the following equivalent form

$$\begin{aligned}\min_{u,p}\quad & \int_\Omega \phi(\kappa_w^k)|p|dx + \frac{r_1}{2}\int_\Omega \left(u - \left(v^{k+1} - \frac{\Lambda_1^k}{r_1}\right)\right)^2 dx\\ \text{s.t.},\quad & p = \nabla_w u.\end{aligned}$$

Likewise, the following augmented Lagrangian functional is considered

$$\mathcal{L}_{\mathrm{NL}}(u,p;\Lambda_2) := \int_\Omega \phi(\kappa_w^k)|p|dx + \frac{r_1}{2}\int_\Omega \left(u - \left(v^{k+1} - \frac{\Lambda_1^k}{r_1}\right)\right)^2 dx + \int_\Omega \Lambda_2(p - \nabla_w u)dx + \frac{r_2}{2}\int_\Omega (p - \nabla_w u)^2 dx,$$

Algorithm 3: The ADMM for solving the minimization u-subproblem

1: **Initialize:** $u^{k+1,0}=u^k$, $p^{k+1,0}=p^k$, $\Lambda_2^{k+1,0}=\Lambda_2^k$;
2: **for** $\ell=0$ to $nInner-1$ **do**
3: Solve $u^{k+1,\ell+1}$ from

$$u^{k+1,\ell+1}=\arg\min_u \frac{r_1}{2}\int_\Omega\Big(u-\Big(v^{k+1}-\frac{\Lambda_1^k}{r_1}\Big)\Big)^2dx+\frac{r_2}{2}\int_\Omega\Big(\nabla_w u-\Big(p^{k,\ell}+\frac{\Lambda_2^{k,\ell}}{r_2}\Big)\Big)^2dx; \tag{24}$$

4: Solve $p^{k+1,\ell+1}$ from

$$p^{k+1,\ell+1}=\arg\min_p \int_\Omega\phi(\kappa_w^k)|p|dx+\frac{r_2}{2}\int_\Omega\Big(p-\Big(\nabla_w u^{k+1,\ell+1}-\frac{\Lambda_2^{k,\ell}}{r_2}\Big)\Big)^2dx; \tag{25}$$

5: Update $\Lambda_2^{k+1,\ell+1}$ from

$$\Lambda_2^{k+1,\ell+1}=\Lambda_2^{k+1,\ell}+r_2(p^{k+1,\ell+1}-\nabla_w u^{k+1,\ell+1});$$

6: **end**
7: **Return:** $u^{k+1}=u^{k+1,nInner}$, $p^{k+1}=p^{k+1,nInner}$, $\Lambda_2^{k+1}=\Lambda_2^{k+1,nInner}$.

which can be efficiently solved by the ADMM-based algorithm; see Algorithm 3.

Now, we are going to discuss the solutions to the sub-minimization problem w.r.t. u and p, both of which can be efficiently solved as follows. For (24), the first-order optimality condition gives the following linear equation

$$\big(r_1\mathcal{I}-r_2\Delta_w\big)u=r_1v^{k+\ell}+\Lambda_1^k-\mathrm{div}_w(r_2p^{k+1,\ell}+\Lambda_2^{k+1,\ell}). \tag{26}$$

Since the graph Laplacian Δ_w is negative semidefinite, $r_1\mathcal{I}-r_2\Delta_w$ is diagonally dominant for the given nonlocal weight function w. Thus, we can compute u using the Gauss-Seidel algorithm similar to [51].

For (25), there is the closed-form solution based on the shrinkage operator, which gives

$$p^{k+1,\ell+1}=\mathrm{shrink}\bigg(\nabla_w u^{k+1,\ell+1}-\frac{\Lambda_2^{k+1,\ell}}{r_2},\frac{\phi(\kappa_w^k)}{r_2}\bigg), \tag{27}$$

where $\mathrm{shrink}(v_1,v_2)=\frac{v_1}{|v_1|}\circ\max\{|v_1|-v_2,0\}$ with the convention $\big(\frac{0}{|0|}=0\big)$ and $\circ$ being the element-wise multiplication.

3.3 Convergence analysis

The convergence analysis of the proximal ADMM in [50] can be utilized to emerge a similar result for our NLκTV method. Since $\phi(\kappa_w^k)$ in (23) corresponds to a constant matrix, the convergence of the sub-minimization problem w.r.t. u follows directly from the classic ADMM such as [21,46]. Then, under the assumption that a solution of (20) exits, we can state the convergence result for Algorithm 2 as follows.

Theorem 2 *Let $\{(v^k,u^k;\Lambda_1^k)\}_{k\in\mathbb{N}}$ be the sequence generated by Algorithm 2 and $(\bar v,\bar u;\bar\Lambda_1)$ be a point satisfying the first-order optimality conditions,*

$$\begin{cases}\lambda K^*(Kv-g)-\Lambda_1=0,\\ -\phi(\kappa_w)\mathrm{div}_w\tau+\Lambda_1=0,\quad \text{where } \tau\in\partial_{\nabla_w u}|\nabla_w u| \text{ and } \phi(\kappa_w)=a+b\Big(\mathrm{div}_w\big(\frac{\nabla_w u}{|\nabla_w u|}\big)\Big)^2,\\ u-v=0.\end{cases} \tag{28}$$

For simplicity, we denote $\partial|\nabla_w u|:=\partial_{\nabla_w u}|\nabla_w u|$. If for any $\tau^k\in\partial|\nabla_w u^k|$ and $\bar\tau\in\partial|\nabla_w\bar u|$ the quantity

$$\Theta_w^k=\Big\langle\big(\phi(\kappa_w^k)-\phi(\bar\kappa_w)\big)\tau^k,\nabla_w u^k-\nabla_w\bar u\Big\rangle,\quad\forall k\in\mathbb{N}, \tag{29}$$

is nonnegative, we have

(1) The successive errors $v^{k+1}-v^k\to 0$, $u^{k+1}-u^k\to 0$, $v^{k+1}-u^k\to 0$ *and* $\Lambda_1^{k+1}-\Lambda_1^k\to 0$ *as* $k\to 0$.
(2) The sequence $(u^k,v^k;\Lambda_1^k)$ *converges to a limit point that satisfies the first-order optimality conditions of* (28).

Proof (1) Let $v_e^k=v^k-\bar{v}$, $u_e^k=u^k-\bar{u}$ and $\Lambda_{1e}^k=\Lambda_1^k-\bar{\Lambda}_1$. It follows from the optimality conditions of v,u,Λ_1 subproblem

$$\begin{cases}\lambda K^*(Kv^{k+1}-g)+r_1(v^{k+1}-v^k)+\gamma(v^{k+1}-v^k)-\Lambda_1^k=0,\\ r_1(u^{k+1}-v^{k+1})-\phi(\kappa_w^{k+1})\mathrm{div}_w\tau^{k+1}+\Lambda_1^k=0,\\ \Lambda_1^{k+1}=\Lambda_1^k+r_1(u^{k+1}-v^{k+1}).\end{cases}\tag{30}$$

In addition, we can express (30) in terms of the error differences v_e,u_e and Λ_{1e} as

$$\begin{cases}(\lambda K^*K+r_1)v_e^{k+1}+\gamma(v_e^{k+1}-v_e^k)=r_1u_e^k+\Lambda_{1e}^k,\\ r_1u_e^{k+1}-(\phi(\kappa_w^{k+1})\mathrm{div}_w\tau^{k+1}-\phi(\bar{\kappa}_w)\mathrm{div}_w\bar{\tau})+\Lambda_{1e}^k=r_1v_e^{k+1},\\ \Lambda_{1e}^{k+1}=\Lambda_{1e}^k+r_1(u_e^{k+1}-v_e^{k+1}).\end{cases}$$

By taking the inner product of the above equations respectively with v_e^{k+1},u_e^{k+1} and Λ_{1e}^k, we have

$$\begin{cases}r_1\|v_e^{k+1}\|^2+\lambda\|v_e^{k+1}\|_K^2+\gamma\langle v_e^{k+1}-v_e^k,v_e^{k+1}\rangle=\langle\Lambda_{1e}^k,v_e^{k+1}\rangle+r_1\langle u_e^k,v_e^{k+1}\rangle,\\ r_1\|u_e^{k+1}\|^2-\langle\phi(\kappa_w^{k+1})\mathrm{div}_w\tau^{k+1}-\phi(\bar{\kappa}_w)\mathrm{div}_w\bar{\tau},u_e^{k+1}\rangle+\langle\Lambda_{1e}^k,u_e^{k+1}\rangle=r_1\langle v_e^{k+1},u_e^{k+1}\rangle,\\ \langle\Lambda_{1e}^{k+1}-\Lambda_{1e}^k,\Lambda_{1e}^k\rangle=r_1\langle u_e^{k+1}-v_e^{k+1},\Lambda_{1e}^k\rangle,\end{cases}$$

with $\|v\|_K^2=\langle Kv,Kv\rangle$. Then, applying $\langle q-q',q\rangle=\frac{1}{2}(\|q\|^2+\|q-q'\|^2-\|q'\|^2)$ to the above equations yeilds

$$\begin{cases}\frac{r_1+\gamma}{2}\|v_e^{k+1}\|^2+\lambda\|v_e^{k+1}\|_K^2+\frac{\gamma}{2}\|v^{k+1}-v^k\|^2+\frac{r_1}{2}\|v_e^{k+1}-u_e^k\|^2=\frac{\gamma}{2}\|v_e^k\|^2+\frac{r_1}{2}\|u_e^k\|^2+\langle\Lambda_{1e}^k,v_e^{k+1}\rangle,\\ \frac{r_1}{2}\|u_e^{k+1}\|^2-\langle\phi(\kappa_w^{k+1})\mathrm{div}_w\tau^{k+1}-\phi(\bar{\kappa}_w)\mathrm{div}_w\bar{\tau},u_e^{k+1}\rangle+\frac{r_1}{2}\|u_e^{k+1}-v_e^{k+1}\|^2+\langle\Lambda_{1e}^k,u_e^{k+1}\rangle=\frac{r_1}{2}\|v_e^{k+1}\|^2,\\ \frac{1}{2r_1}\|\Lambda_{1e}^{k+1}\|^2=\frac{1}{2r_1}\|\Lambda_{1e}^k\|^2+\frac{r_1}{2}\|u_e^{k+1}-v_e^{k+1}\|^2+\langle u_e^{k+1}-v_e^{k+1},\Lambda_{1e}^k\rangle.\end{cases}\tag{31}$$

Denote

$$e^k:=\frac{r_1}{2}\|u_e^k\|^2+\frac{\gamma}{2}\|v_e^k\|^2+\frac{1}{2r_1}\|\Lambda_{1e}^k\|^2.$$

Then, summing of the equations in (31), we obtain

$$e^{k+1}+\frac{r_1}{2}\|v_e^{k+1}-u_e^k\|^2+\frac{\gamma}{2}\|v^{k+1}-v^k\|^2+\lambda\|v_e^{k+1}\|_K^2-\langle\phi(\kappa_w^{k+1})\mathrm{div}_w\tau^{k+1}-\phi(\bar{\kappa}_w)\mathrm{div}_w\bar{\tau},u_e^{k+1}\rangle=e^k,$$

where $\|v_e^{k+1}\|_K^2\geq 0$ and $-\langle\phi(\kappa_w^{k+1})\mathrm{div}_w\tau^{k+1}-\phi(\bar{\kappa}_w)\mathrm{div}_w\bar{\tau},u_e^{k+1}\rangle=\langle\phi(\kappa_w^{k+1})\tau^{k+1}-\phi(\bar{\kappa}_w)\bar{\tau},\nabla_w u_e^{k+1}\rangle=\Theta_w^{k+1}+\phi(\bar{\kappa})\langle\tau_e^{k+1},\nabla_w u_e^{k+1}\rangle$. Referring to Lemma 3.3 in [13], the term $\langle\tau_e^{k+1},\nabla_w u_e^{k+1}\rangle\geq 0$ for any $\tau^k\in\partial|\nabla_w u^k|$ and $\bar{\tau}\in\partial|\nabla_w\bar{u}|$. Based on the assumption on $\Theta_w^k\geq 0$ for all k, hence $-\langle\phi(\kappa_w^{k+1})\mathrm{div}_w\tau^{k+1}-\phi(\bar{\kappa}_w)\mathrm{div}_w\bar{\tau},u_e^{k+1}\rangle\geq 0$. By dropping the two nonnegtive terms, we have

$$e^{k+1}+\frac{r_1}{2}\|v_e^{k+1}-u_e^k\|^2+\frac{\gamma}{2}\|v^{k+1}-v^k\|^2\leq e^k.\tag{32}$$

Then $\{e^k\}_{k\in\mathbb{N}}$ is a monotone decreasing sequence. Therefore, by further summing (32) from $k=0$ to ∞, it implies that

$$\frac{r_1}{2}\sum_{k=0}^{\infty}\|v_e^{k+1}-u_e^k\|^2+\frac{\gamma}{2}\sum_{k=0}^{\infty}\|v^{k+1}-v^k\|^2\leq\sum_{k=0}^{\infty}e^k-e^{k+1}<\infty.$$

Thus, there is $v^{k+1}-v^k\to 0$ and $v^{k+1}-u^k\to 0$ as $k\to\infty$. Besides, by applying Minkowski's inequality, we also obtain

$$\begin{cases}\lim\limits_{k\to\infty}\|u^{k+1}-u^k\|\le\lim\limits_{k\to\infty}\left(\|u_e^{k+1}-v_e^{k+2}\|+\|v^{k+2}-v^{k+1}\|+\|v_e^{k+1}-u_e^k\|\right)=0,\\ \lim\limits_{k\to\infty}\|\Lambda_1^{k+1}-\Lambda_1^k\|\le\lim\limits_{k\to\infty}\left(r_1\|u_e^{k+1}-v_e^{k+2}\|+r_1\|v_e^{k+2}-v_e^{k+1}\|\right)=0.\end{cases}$$

That is, $u^{k+1}-u^k\to 0$ and $\Lambda_1^{k+1}-\Lambda_1^k\to 0$ as $k\to\infty$.
(2) We showed the error $\{e_k\}_{k\in\mathbb{N}}$ is a monotone decreasing sequence in $\mathbb{R}^+$ for any point $(\bar v,\bar u;\bar\Lambda_1)$. Thus, the sequence $\{(v^k,u^k;\Lambda_1^k)\}_k$ is uniformly bounded on Ω. Hence, a weakly convergent subsequence $\{(v^{k_\ell},u^{k_\ell};\Lambda_1^{k_\ell})\}_{\ell\in\mathbb{N}}$ exists and has the limit point $(v^*,u^*;\Lambda_1^*)$. Since $u^{k_\ell}\to u^*$ a.e. in Ω as $\ell\to\infty$, $\nabla_w u^{k_\ell}\to\nabla_w u^*$ a.e. in Ω as $\ell\to\infty$. According to Lemma 1 in [50], there exists a subsequence of $\{\tau^{k_\ell}\}_{\ell\in\mathbb{N}}$ that converges weakly to $\tau^*\in\partial|\nabla_w u^*|$.

The sequence $\{(v^{k_\ell},u^{k_\ell};\Lambda_1^{k_\ell})\}_{\ell\in\mathbb{N}}$ satisfies in the optimality conditions (30), i.e.,

$$\begin{cases}\lambda K^*(Kv^{k_\ell+1}-g)+r_1(v^{k_\ell+1}-v^{k_\ell})+\gamma(v^{k_\ell+1}-v^{k_\ell})-\Lambda_1^{k_\ell}=0,\\ r_1(u^{k_\ell+1}-v^{k_\ell+1})-\phi(\kappa_w^{k_\ell+1})\mathrm{div}_w\tau^{k_\ell+1}+\Lambda_1^{k_\ell}=0,\\ \Lambda_1^{k_\ell+1}=\Lambda_1^{k_\ell}+r_1(u^{k_\ell+1}-v^{k_\ell+1}).\end{cases}$$

Based on part (1), taking the limit from the convergent subsequence, we obtain

$$\begin{cases}\lambda K^*(Kv^*-g)-\Lambda_1^*=0,\\ -\phi(\kappa_w^*)\mathrm{div}_w\tau^*+\Lambda_1^*=0,\ \tau\in\partial|\nabla_w u^*|,\\ u^*-v^*=0,\end{cases}$$

for almost every point in Ω. This derives that the limit point satisfies the first-order optimality conditions (28). This completes the proof of the convergence theorem. □

Remark 2 Similarly to the κTV method, we introduce the proximal term of v to achieve the theoretical convergence of Algorithm 2, and set $\gamma=0$ in the numerical experiments. We also observe that the Θ_w^k can be guaranteed nonnegative numerically as shown in Fig. 7.

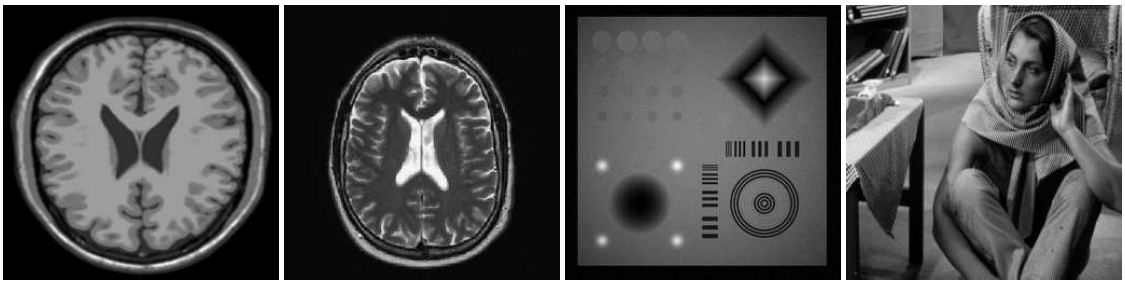
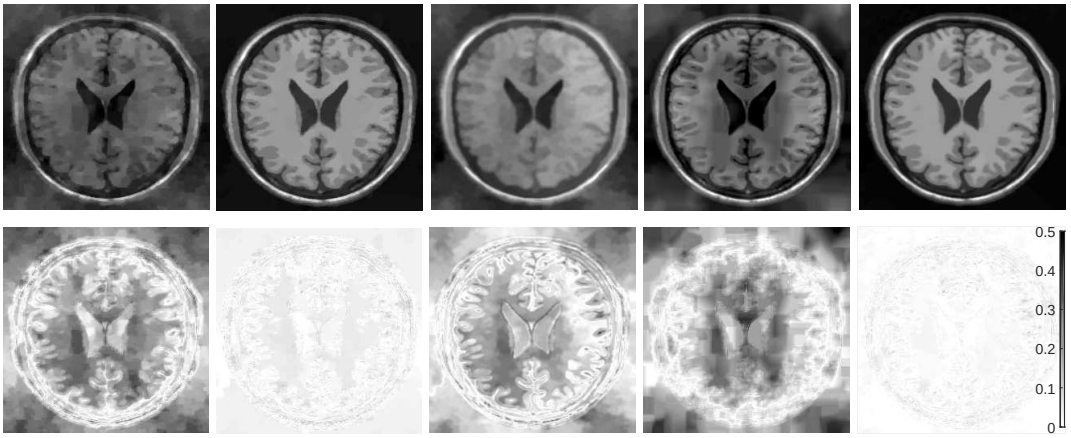


Fig. 2 The original test images. From left to right: T_1-weighted brain, T_2-weighted brain, Phantom and Barbara

4 Numerical experiments

To assess the performance of the Euler's elastica model for sparse reconstruction, we carry out simulations on one synthetic T_1-weighted brain data[1], one T_2-weighted brain data[2], a phantom image with

[1] http://www.bic.mni.mcgill.ca/brainweb/.
[2] http://www.mr-tip.



(a) TV (b) κTV (c) NLTV (d) BM3D (e) NLκTV

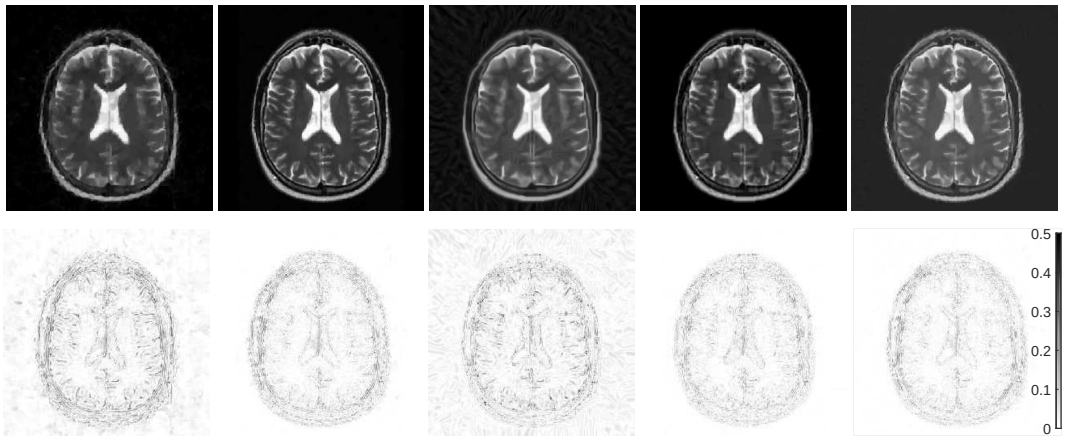
Fig. 3 The reconstructed images (the first row) and the error images $|u - u^*|$ (the second row) of the T_1-weighted brain image under 10% random sampling pattern

Table 1 The PSNR performance comparison among MR reconstruction approaches for different sampling masks and sampling ratio of 10%

Image	T1w Brain			T2w Brain			Phantom			Barbara		
	Ran.	Cart.	Radial	Ran.	Cart.	Radial	Ran.	Cart.	Radial	Ran.	Cart.	Radial
Zero-filling	13.37	19.69	22.11	14.94	19.73	21.18	17.18	22.70	26.75	16.84	19.97	22.37
TV	14.77	21.13	27.38	23.95	21.49	26.03	22.85	25.53	30.46	19.47	21.50	24.70
TGVST	15.69	24.14	31.53	18.88	23.40	28.90	18.54	24.99	32.61	18.39	21.71	25.26
κTV	30.01	25.02	31.84	31.99	23.55	28.63	23.99	28.17	31.66	21.46	22.92	25.87
LaSAL2	13.68	22.26	30.28	14.68	22.57	28.63	18.22	24.33	30.03	17.76	21.45	24.52
WaTMRI	15.35	22.95	31.10	22.26	22.68	28.49	19.63	25.34	30.60	19.01	21.85	24.46
NLTV	17.68	25.07	30.37	24.79	23.98	27.69	24.74	28.63	34.66	21.56	23.24	26.07
PANO	16.69	24.67	31.94	18.54	23.14	28.45	22.07	26.21	35.44	20.42	22.83	26.33
BM3D	15.80	24.72	33.47	16.89	23.87	29.74	19.24	27.91	35.61	19.72	22.27	26.46
NLκTV	**33.22**	**25.75**	**34.07**	**33.37**	**24.02**	**29.80**	**24.84**	**29.23**	**35.93**	**22.05**	**23.43**	**26.53**

homogeneous structures and the Barbara image with abundant texture patterns; see Fig. 2. The data acquisition is simulated by undersampling the 2D discrete Fourier transform coefficients of the test images using three different patterns, i.e., random, Cartesian and radial pattern. All images are normalized to [0,1] and the sampling ratio is chosen from 10% to 50%.

We compare the proposed algorithms with the state-of-the-art CS-MRI methods including zero-filling [6], TV [32,24], NLTV [51], TGVST [22], WaTMRI [12], LaSAL2 [38], PANO [40] and BM3D method [17]. The TV and TGVST are two comparable approaches to the κTV method, the regularization of which involves the TV term. The NLTV, PANO and BM3D method, which make use of the powerful nonlocal similarities for reconstruction, are mostly relevant to our NLκTV method. Besides, WaTMRI and LaSAL2 are two representatives of wavelet-based approaches by exploiting the wavelet tree structure for sparse reconstruction. All experiments were performed in MATLAB on a dell desktop with 3.6GHz Intel Core i7.



(a) TV (b) κTV (c) NLTV (d) BM3D (e) NLκTV

Fig. 4 The reconstructed images (the first row) and the error images $|u - u^*|$ (the second row) of the T_2-weighted brain image under 10% radial sampling pattern

Table 2 The PSNR evaluations among different MR reconstruction approaches for different sampling masks and sampling ratio of 20%

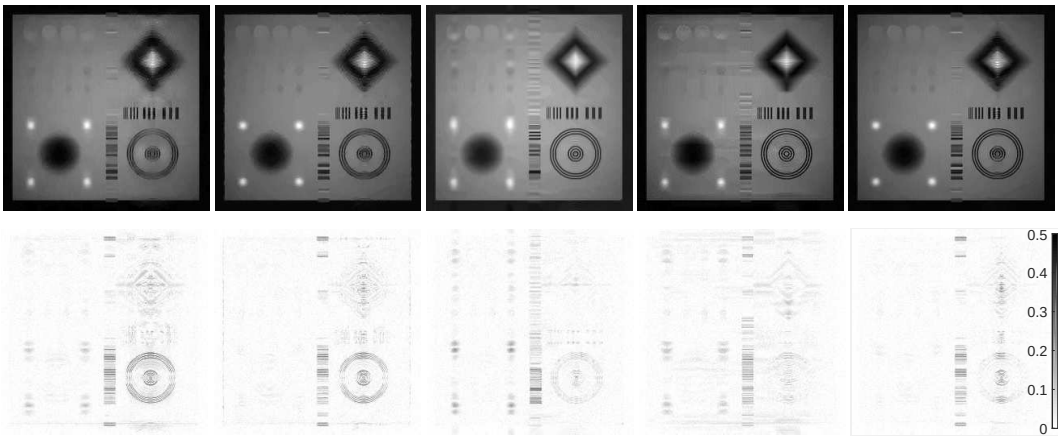
Image	T1w Brain			T2w Brain			Phantom			Barbara		
	Ran.	Cart.	Radial	Ran.	Cart.	Radial	Ran.	Cart.	Radial	Ran.	Cart.	Radial
Zero-filling	20.07	21.84	28.32	21.27	21.40	24.65	33.02	26.69	29.22	21.33	22.90	25.98
TV	33.54	24.03	36.99	34.39	23.26	33.80	33.50	29.79	35.46	27.10	24.55	28.82
TGVST	41.81	25.38	39.78	36.62	24.36	35.41	36.97	30.15	35.99	29.23	24.91	30.06
κTV	41.58	26.35	39.99	36.75	24.85	36.19	36.74	31.17	36.05	30.22	25.56	29.81
LaSAL2	40.41	25.40	41.67	37.87	24.25	37.69	37.40	30.56	35.25	27.39	23.98	27.57
WaTMRI	41.02	24.93	39.30	36.31	23.87	35.92	37.37	29.79	36.48	27.86	24.65	29.12
NLTV	39.41	26.79	38.08	34.80	25.18	33.22	38.71	32.22	38.07	31.22	26.36	31.04
PANO	44.34	26.59	42.60	37.66	24.71	36.88	38.76	30.07	36.47	31.16	26.20	30.33
BM3D	44.96	**27.10**	44.29	**38.18**	25.40	37.50	39.01	32.37	38.51	**31.63**	25.92	**31.34**
NLκTV	**45.40**	26.82	**44.81**	38.14	**25.60**	**38.33**	**39.02**	**33.05**	**39.23**	31.42	**26.43**	31.18

4.1 Optimal parameters settings

There are two important parameters for the Euler's elastica regularity, i.e., a and b, which are fixed as $a = 8, b = 10$ for the κTV and $a = 0.01$, $b = 0.01$ for the NLκTV, respectively. And ς is set as $\varsigma = 10^{-6}$ in both κTV and NLκTV model. We fix $nIter = 700$ and select λ and r in between the range $[5\times10^3, 3\times10^4]$ for the κTV. On the other hand, for NLκTV, we empirically set $\lambda = 10^4$, $nIter = 100$, $nInner = 2$ and select $r_1 \in [10, 100]$, $r_2 \in [0.1, 1]$ in most experiments.

For both NLTV and NLκTV method, $\forall x \in \Omega$, we use the best 10 neighbors in the semi-local searching window of size 11×11 centered at x with the patch size of 5×5. Other parameters for the NLTV method are selected as $\lambda_{\text{NLTV}} = 0.5$, $r \in [0.001, 0.01]$, $nIter = 500$ and the inner iteration is set as $nInner = 10$ for the denoising step.

Parameters for the comparison algorithms are chosen to achieve the best balance of measure and CPU-time in each experiment. The regularization parameter λ plays an important role in the TV and PANO model, which is chosen as $\lambda_{\text{TV}} \in [10^2, 10^4]$ and $\lambda_{\text{PANO}} = 10^6$ in most experiments. For BM3D, the range of the parameter for the observation fidelity is $\lambda_{\text{BM3D}} \in [0, 5]$. The total iteration number for the TGVST, LaSAL2 and BM3D is set as 500, 100 and 80 to balance the performance and efficiency.



(a) TV (b) κTV (c) NLTV (d) BM3D (e) NLκTV

Fig. 5 The reconstructed images (the first row) and the error images $|u - u^*|$ (the second row) of the phantom data under 20% Cartesian sampling pattern

Table 3 The performance evaluations among MR reconstruction approaches for different sampling masks and sampling ratio of 50%

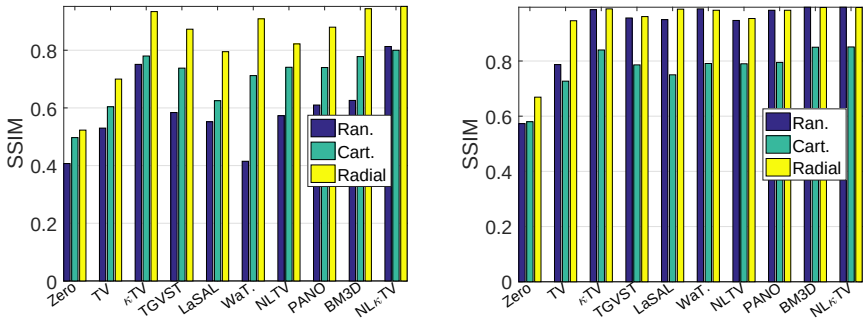
Image	T1w Brain						T2w Brain					
	Random		Cartesian		Radial		Random		Cartesian		Radial	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Zero-filling	42.33	0.9098	30.84	0.8067	40.51	0.8779	33.72	0.7138	28.34	0.8196	32.74	0.6862
TV	50.21	0.9985	38.23	0.9684	50.23	0.9985	44.59	0.9949	35.44	0.9643	45.59	0.9951
TGVST	51.26	0.9952	39.76	0.9702	50.64	0.9932	45.81	0.9872	38.16	0.9739	46.23	0.9845
κTV	52.04	0.9987	44.43	0.9943	52.02	0.9987	45.19	0.9954	39.80	0.9848	46.28	0.9957
LaSAL2	53.85	0.9986	43.07	0.9887	54.56	0.9987	47.75	0.9936	38.69	0.9736	49.22	0.9948
WaTMRI	54.10	0.9992	41.44	0.9882	**55.47**	**0.9994**	49.56	**0.9975**	39.64	0.9819	**51.11**	**0.9980**
NLTV	51.91	0.9974	40.30	0.9598	51.64	0.9972	42.68	0.9701	38.23	0.9605	43.45	0.9741
PANO	**54.83**	0.9992	42.16	0.9824	55.25	0.9992	49.56	0.9972	39.95	0.9807	50.78	0.9974
BM3D	54.57	0.9992	47.77	**0.9972**	55.18	0.9992	**49.58**	0.9962	**42.68**	**0.9893**	50.86	0.9964
NLκTV	**54.83**	**0.9993**	**47.98**	0.9967	55.42	0.9993	49.40	0.9970	42.36	**0.9893**	50.76	0.9975

Both the iteration number and relative error $\epsilon = 5 \times 10^{-5}$ are used as the terminating conditions for all algorithms. We evaluate the reconstruction performance using both PSNR (peak signal-to-noise ratio) and SSIM (structural similarity index).

4.2 The performance evaluations

In the first example, we generate the undersampled k-space data with sampling ratio 10% and three kinds of sampling patterns, where Table 1 details the PSNR obtained by the comparison algorithms. It is obviously shown that the NLκTV method outperforms other CS-MRI approaches. We display both the reconstructed image and error image of T_1-weighted brain image using random sampling and the T_2-weighted brain image using radial sampling in Fig. 3 and Fig. 4, respectively. Based on the qualitative and quantitative results, we make the following observations:

① The κTV produces much better reconstruction results than TV model, which demonstrates that the minimization of the total curvature can help to recover the missing information and effectively improve the reconstruction quality;



(a) Sampling ratio = 10% (b) Sampling ratio = 20%

Fig. 6 The SSIM evaluations of T_1-weighted brain image for different sampling masks and sampling ratios

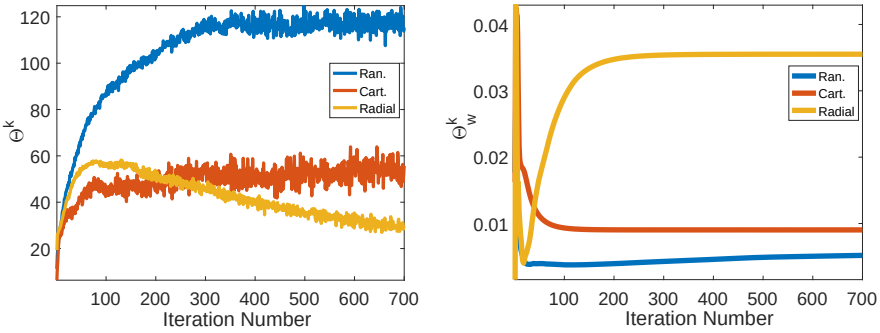
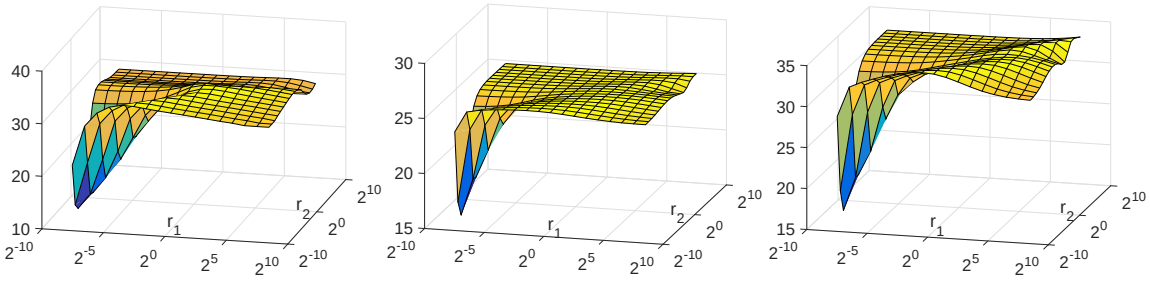


Fig. 7 On the T_1-weighted image with different sampling masks, the behavior of Θ^k (left) and Θ_w^k (right) versus iteration numbers

② It is well-known that the nonlocal approaches can take care of the texture structures (e.g., Barbara image) based on patch redundancy. Indeed, the comparisons between NLTV and NLκTV show that the nonlocal curvature not only preserves the fine structures and details but also the homogeneity in the smooth region such as the white matter of the T_1-weighted brain image and black background of the T_2-weighted brain image;

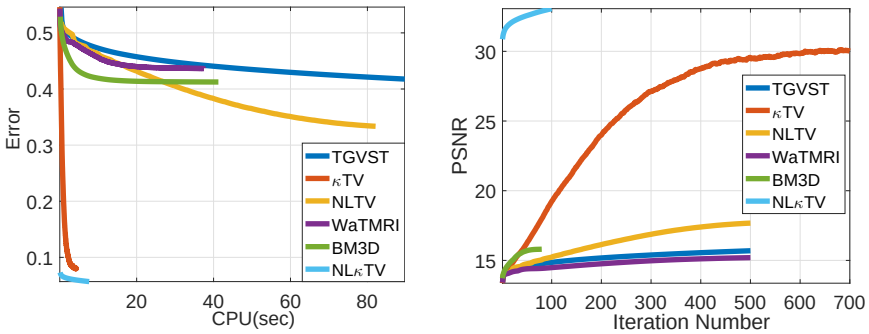
③ By comparing the κTV and NLκTV, especially the error images, we observe the NLκTV always presents a lighter coloring, which indicates the reconstruction of NLκTV is much closer to the ground truth. In addition, the PSNR values also prove the advantages of the nonlocal similarities in improving the reconstruction quality.

Table 2 lists the PSNR obtained by the CS-MRI algorithms on undersampled data with sampling ratio 20% and different sampling patterns. Clearly, the performances of the proposed methods are consistent with the experiments using sampling ratio 10%, such that our approaches give better reconstruction results for most image- and mask-type combinations. As shown, the PSNR values of NLκTV exceed the ones obtained by NLTV and κTV in all experiments. Especially, the NLκTV is noticeably better than all local methods because more patch similarities can be observed with relatively high sampling rate. We also display the selective visual results of the phantom image with Cartesian sampling pattern in Fig. 5. Compared to the NLTV and BM3D approach, less energy is contained in the left part of the error image of our NLκTV method. However, the upper right corner of NLκTV is not as smooth as NLTV because the weight function is precomputed based on the κTV reconstruction and fixed during the reconstruction. Even though, the overall quality of the NLκTV method is better than NLTV and BM3D for the phantom image.



(a) Random sampling (b) Cartesian sampling (c) Radial sampling

Fig. 8 Reconstruction PSNR in dB of the NLκTV with different r_1, r_2 for 10% random, Cartesian and radial sampling pattern



(a) Absolute error versus CPU-time (b) PSNR versus iteration numbers

Fig. 9 Evaluations on the T_1-weighted image with 10% random downsampling

Fig. 6 depicts the SSIM values evaluated on the T_1-weighted brain image for three mask types with sampling ratio 10% and 20%, which correspond with the PSNR listed in Table 1 and 2. Similarly to the denoising example in Fig. 1, the nonlocal elastica model significantly improves the image structural information compared to the NLTV method. Among these CS-MRI algorithms, the nonlocal methods (i.e., NLTV, PANO and BM3D) perform much better than other approaches, especially the BM3D which utilizes the Block Matching 3D (BM3D) model as the regularizer for MRI reconstruction problem. By facilitating nonlocal structures in the image using groups of image patches, the BM3D model can achieve the state-of-the-art performance for image denoising and reconstruction problems. We observe that our NLκTV model can give better (e.g. the reconstruction image with random sampling and sampling rate 10%) or comparable reconstruction results than the BM3D model due to the minimization of the nonlocal curvature.

In addition, we carry out an experiment on a relatively high sampling rate 50% to make the comparison more convincing. We use the T_1-weighted and T_2-weighted brain image as examples and record both the PSNR and SSIM in Table 3. Compared to sampling rate of 20%, the performance of all CS-MRI reconstruction methods increase significantly. The WaTMRI and PANO work well on the random and radial pattern, but their performance on Cartesian pattern are not as good as BM3D and NLκTV. As shown, our NLκTV model provides very competitive results for all sampling patterns, which has a remarkable advantage over both κTV and NLTV.

To validate the numerical convergence of the proposed algorithms, we plot the curves of Θ^k and Θ_w^k using the T_1-weighted image of sampling ratio 10% in Fig. 7. In the computation of Θ^k and Θ_w^k, τ^k is the subgradient of L^1 norm, which is simply set to zero at the non-differentiable points. Although we can

not find the theoretical lower bound of Θ^k and Θ_w^k, it is shown to be nonnegative in practice. Besides, we discuss the impact of r_1 and r_2 in NLκTV method based on the same test image. By varying the parameter $(r_1, r_2) \in [2^{-10}, 2^{10}] \times [2^{-10}, 2^{10}]$, we perform the NL$\kappa$TV algorithm and plot the PSNR in Fig. 8. It can be observed that there are relatively large intervals for both r_1 and r_2 to generate good reconstruction results.

Fig. 9 presents the absolute error and PSNR curve obtained using the T_1-weighted brain image under the random sampling pattern. As we mentioned, the weight function of our NLκTV method is computed based on the κTV reconstruction. The error curve shows that both κTV and NLκTV method converge very fast, i.e., less than 10 seconds, while other methods require more than 30 seconds and converge to relatively large absolute errors. On the other hand, the PSNR curves illustrate that both NLκTV and BM3D method converge less than 100 iterations and our NLκTV method achieves results with much better quality. Thus, our NLκTV algorithm outperforms other CS-MRI methods in numerical convergence.

4.3 Estimation of the weight function

The weight function plays a very critical role in nonlocal methods, especially for image inpainting and sparse reconstruction problem. We compare the reconstruction quality and the computational efficiency of the NLκTV method with either updated or fixed weights on the T_1-weighted brain image with sampling ratio 20%. The terminating conditions are given as either the relative error reaching 3×10^{-5} or the iteration number reaching 500. In particular, the initial weight for the updated case is calculated using the zero-filling reconstruction and updated every 20 iterations thereafter. We list the obtained PSNR and SSIM in Table 4 which shows the NLκTV with the fixed weights can provide reconstruction results as good as the ones obtained using the updated weights. On the other hand, the CPU-time of the fixed weight scheme is far faster. Thus, it is reasonable to use the reconstruction of the κTV method to compute the weight function for the NLκTV model.

Table 4 The performance evaluation and speed comparison for fixed weights and updated weights on the T_1-weighted image of sampling ratio 20%

Mask	Weights	PSNR	SSIM	Iteration	CPU(sec)
Random	updated	45.52	0.993	320	96.0
	fixed	45.41	0.995	54	8.7
Cartesian	updated	26.59	0.888	500	144.0
	fixed	26.82	0.851	44	7.8
Radial	updated	45.03	0.994	117	34.4
	fixed	44.85	0.994	73	9.9

4.4 Complexity analysis

The proposed NLκTV method can be separated into the reconstruction step and nonlocal denoising step. The reconstruction (22) is realized by the FFT, the computational complexity of which is $\mathcal{O}(N \log N)$ with N being the total pixels of an image. The computational complexity of nonlocal denoising is $\mathcal{O}(NQ)$, where Q is the number of similar patches used in the experiments. Without loss of generality, the computational complexity of the NLκTV method can be approximately given as $\mathcal{O}(N \log N)$; see Table 5.

We also list the average time required by each algorithm obtained by averaging among all test images with all combinations of the mask types and sampling ratios in Table 5, where the CPU-time of NLκTV

is the total time of the preprocessing using the κTV and the reconstruction by nonlocal κTV. As shown, although the computational complexity is the same as the NLTV and PANO, the NLκTV costs much less CPU-time as it converges faster in practice. Besides, we observe that our NLκTV is also faster than the BM3D such that our proposal produces comparable or better reconstruction results using only less than 1/3 its CPU-time.

Table 5 The complexity and average CPU-time for different MR reconstruction methods

Methods	κTV	TGVST	LaSAL2	WaTMRI	NLTV	PANO	BM3D	NLκTV
$\mathcal{O}(\cdot)$	$N \log N$	$N \log N$	-	$N \log N$	$N \log N$	NQ	$N \log N$	$N \log N$
Time(s)	4.5	34.3	56.9	5.5	81.3	57.4	33.5	10.2

5 Conclusion

We proposed a novel sparse reconstruction method based on the Euler's elastica energy, which is a well-known high-order regularizer for dealing with image denoising and inpainting problem. In order to further improve the reconstruction quality, we extended the Euler's elastica into the nonlocal formulation. By regarding the Euler's elastica as a weighted TV model, efficient algorithms were developed for both local and nonlocal model based on the ADMM, the convergence of which was also discussed under certain assumptions. We evaluated both local and nonlocal elastica methods on sparse reconstruction problems under different sampling patterns and sampling ratios for a variety of images. By compared with the state-of-the-art CS-MRI algorithms, our proposal produced the overall best performance for undersampled observations with high efficiency. It demonstrated the superiority of the NLκTV method in recovering missing information and preserving structures by taking advantages of both high-order regularizer and the nonlocal mechanism.

The proposed nonlocal elastica model can be applied to many other image tasks such as denoising, inpainting and segmentation, in which the minimization of curvature can play an important role. We can also further improve our nonlocal formulation using the block based method [30] and measuring the distances between image patches by structural similarity [45].

Acknowledgements The work is supported by NSFC 11701418, Major Science and Technology Project of Tianjin 18ZXRHSY00160 and Recruitment Program of Global Young Expert.

References

1. Ambrosio, L., Masnou, S.: On a variational problem arising in image reconstruction. Isnm International **147**, 17–26 (2006)
2. Arias, P., Caselles, V., Sapiro, G.: A variational framework for non-local image inpainting. In: International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition, pp. 345–358 (2009)
3. Arias, P., Facciolo, G., Caselles, V., Sapiro, G.: A variational framework for exemplar-based image inpainting. International Journal of Computer Vision **93**(3), 319–347 (2011)
4. Bae, E., Shi, J., Tai, X.C.: Graph cuts for curvature based image denoising. IEEE Transactions on Image Processing **20**(5), 1199–1210 (2011)
5. Bae, E., Tai, X.C., Zhu, W.: Augmented Lagrangian method for an Euler's elastica based segmentation model that promotes convex contours. Inverse Problems & Imaging **11**(1), 1–23 (2017)
6. Bernstein, M.A., Fain, S.B., Riederer, S.J.: Effect of windowing and zero-filled reconstruction of MRI data on spatial resolution and acquisition strategy. Journal of Magnetic Resonance Imaging **14**(3), 270–280 (2001)
7. Bredies, K., Kunisch, K., Pock, T.: Total generalized variation. SIAM Journal on Imaging Sciences **3**(3), 492–526 (2010)
8. Buades, A., Coll, B., Morel, J.M.: A review of image denoising algorithms, with a new one. Multiscale Modeling & Simulation **4**(2), 490–530 (2005)

9. Chambolle, A.: An algorithm for total variation minimization and applications. Journal of Mathematical Imaging and Vision **20**(1-2), 89–97 (2004)
10. Chambolle, A., Pock, T.: A first-order primal-dual algorithm for convex problems with applications to imaging. Journal of Mathematical Imaging and Vision **40**(1), 120–145 (2011)
11. Chan, T.F., Kang, S.H., Shen, J.: Euler's elastica and curvature-based inpainting. SIAM Journal on Applied Mathematics **63**(2), 564–592 (2002)
12. Chen, C., Huang, J.: Exploiting the wavelet structure in compressed sensing MRI. Magnetic Resonance Imaging **32**(10), 1377–1389 (2014)
13. Chen, Y., Hager, W.W., Yashtini, M., Ye, X., Zhang, H.: Bregman operator splitting with variable stepsize for total variation image reconstruction. Computational Optimization and Applications **54**(2), 317–342 (2013)
14. Compton, R., Osher, S., Bouchard, L.S.: Hybrid regularization for MRI reconstruction with static field inhomogeneity correction. Inverse Problems & Imaging **7**(4), 1215–1233 (2013)
15. Deng, L.J., Glowinski, R., Tai, X.C.: A new operator splitting method for the Euler elastica model for image smoothing. SIAM Journal on Imaging Sciences **12**(2), 1190–1230 (2019)
16. Duan, Y., Huang, W., Zhou, J., Chang, H., Zeng, T.: A two-stage image segmentation method using Euler's elastica regularized Mumford-Shah model. In: 2014 22nd International Conference on Pattern Recognition, pp. 118–123 (2014)
17. Eksioglu, E.M.: Decoupled algorithm for MRI reconstruction using nonlocal block matching model: BM3D-MRI. Journal of Mathematical Imaging and Vision **56**(3), 430–440 (2016)
18. Esedoglu, S., March, R.: Segmentation with depth but without detecting junctions. Journal of Mathematical Imaging & Vision **18**(1), 7–15 (2003)
19. Gilboa, G., Osher, S.: Nonlocal linear image regularization and supervised segmentation. Multiscale Modeling & Simulation **6**(2), 595–630 (2007)
20. Gilboa, G., Osher, S.: Nonlocal operators with applications to image processing. SIAM Journal on Multiscale Modeling & Simulation **7**(3), 1005–1028 (2007)
21. Goldstein, T., Osher, S.: The split Bregman method for L_1 regularized problems. SIAM Journal on Imaging Sciences **2**(2), 323–343 (2009)
22. Guo, W., Qin, J., Yin, W.: A new detail-preserving regularization scheme. SIAM Journal on Imaging Sciences **7**(2), 1309–1334 (2014)
23. Guo, W., Yin, W.: Edge guided reconstruction for compressive imaging. SIAM Journal on Imaging Sciences **5**(3), 809–834 (2012)
24. Haldar, J.P., Diego, H., Zhi-Pei, L.: Compressed-sensing MRI with random encoding. IEEE Transactions on Medical Imaging **30**(4), 893–903 (2011)
25. Hammernik, K., Klatzer, T., Kobler, E., Recht, M.P., Sodickson, D.K., Pock, T., Knoll, F.: Learning a variational network for reconstruction of accelerated MRI data. Magnetic Resonance in Medicine **79**(6), 3055–3071 (2018)
26. He, X., Zhu, W., Tai, X.C.: Segmentation by elastica energy with L_1 and L_2 curvatures: a performance comparison. Numerical Mathematics-Theory Methods and Applications **12**(1), 285–311 (2019)
27. Huang, J., Yang, F.: Compressed magnetic resonance imaging based on wavelet sparsity and nonlocal total variation. In: 2012 9th IEEE International Symposium on Biomedical Imaging (ISBI), pp. 968–971 (2012)
28. Kang, S.H., Zhu, W., Jianhong, J.: Illusory shapes via corner fusion. SIAM Journal on Imaging Sciences **7**(4), 1907–1936 (2014)
29. Knoll, F., Bredies, K., Pock, T., Stollberger, R.: Second order total generalized variation (TGV) for MRI. Magnetic Resonance in Medicine **65**(2), 480–491 (2011)
30. Liu, J., Zheng, X.: A block nonlocal TV method for image restoration. SIAM Journal on Imaging Sciences **10**(2), 920–941 (2017)
31. Lou, Y., Zhang, X., Osher, S., Bertozzi, A.: Image recovery via nonlocal operators. Journal of Scientific Computing **42**(2), 185–197 (2010)
32. Lustig, M., Donoho, D., Pauly, J.M.: Sparse MRI: The application of compressed sensing for rapid MR imaging. Magnetic Resonance in Medicine **58**(6), 1182–1195 (2007)
33. Lustig, M., Donoho, D.L., Santos, J.M., Pauly, J.M.: Compressed sensing MRI. IEEE Signal Processing Magazine **25**(2), 72–82 (2008)
34. Masnou, S., Morel, J.M.: Level lines based disocclusion. In: Proceedings 1998 International Conference on Image Processing, pp. 259–263 (1998)
35. Masnou, S., Morel, J.M.: On a variational theory of image amodal completion. Rendiconti Del Seminario Matematico Della Universita Di Padova **116**(4), 211–252 (2005)
36. Mumford, D.: Elastica and computer vision. In: Algebraic Geometry and its Applications, pp. 491–506 (1994)
37. Nesterov, Y.E.: A method for solving the convex programming problem with convergence rate $O(1/k^2)$. Dokl. Akad. Nauk SSSR **269**(3), 543–547 (1983)
38. Panić, M., Aelterman, J., Crnojević, V., Pižurica, A.: Sparse recovery in magnetic resonance imaging with a Markov random field prior. IEEE Transactions on Medical Imaging **36**(10), 2104–2115 (2017)
39. Qu, X., Guo, D., Ning, B., Hou, Y., Lin, Y., Cai, S., Chen, Z.: Undersampled MRI reconstruction with patch-based directional wavelets. Magnetic Resonance Imaging **30**(7), 964–977 (2012)
40. Qu, X., Hou, Y., Lam, F., Guo, D., Zhong, J., Chen, Z.: Magnetic resonance image reconstruction from undersampled measurements using a patch-based nonlocal operator. Medical Image Analysis **18**(6), 843–856 (2014)

41. Ravishankar, S., Bresler, Y.: MR image reconstruction from highly undersampled k-space data by dictionary learning. IEEE Transactions on Medical Imaging **30**(5), 1028–1041 (2011)
42. Romberg, J.: Imaging via compressive sampling. IEEE Signal Processing Magazine **25**(2), 14–20 (2008)
43. Tai, X.C., Hahn, J., Chung, G.J.: A fast algorithm for Euler's elastica model using augmented Lagrangian method. SIAM Journal on Imaging Sciences **4**(1), 313–344 (2011)
44. Tomasi, C., Manduchi, R.: Bilateral filtering for gray and color images. In: International Conference on Computer Vision, pp. 839–846 (1998)
45. Wang, W., Li, F., Ng, M.K.: Structural similarity based nonlocal variational models for image restoration. IEEE Transactions on Image Processing (2019)
46. Wu, C., Tai, X.C.: Augmented Lagrangian method, dual methods, and split Bregman iteration for ROF, vectorial TV, and high order models. SIAM Journal on Imaging Sciences **3**(3), 300–339 (2010)
47. Yang, G., Yu, S., Dong, H., Slabaugh, G., Dragotti, P.L., Ye, X., Liu, F., Arridge, S., Keegan, J., Guo, Y.: Dagan: deep de-aliasing generative adversarial networks for fast compressed sensing MRI reconstruction. IEEE Transactions on Medical Imaging **37**(6), 1310–1321 (2018)
48. Yang, Y., Sun, J., Li, H., Xu, Z.: Deep ADMM-Net for compressive sensing MRI. Advances in Neural Information Processing Systems pp. 10–18 (2016)
49. Yaroslavsky, L.P.: Digital picture processing: an introduction. Springer Ser. Inform. Sci. (1985)
50. Yashtini, M., Kang, S.H.: A fast relaxed normal two split method and an effective weighted TV approach for Euler's elastica image inpainting. SIAM Journal on Imaging Sciences **9**(4), 1552–1581 (2016)
51. Zhang, X., Burger, M., Bresson, X., Osher, S.: Bregmanized nonlocal regularization for deconvolution and sparse reconstruction. SIAM Journal on Imaging Sciences **3**(3), 253–276 (2010)
52. Zhili, Y., Mathews, J.: Nonlocal regularization of inverse problems: a unified variational framework. IEEE Transactions on Image Processing **22**(8), 3192–3203 (2013)
53. Zhu, W., Chan, T., Esedoglu, S.: Segmentation with depth: a level set approach. SIAM Journal on Scientific Computing **28**(5), 1957–1973 (2006)
54. Zhu, W., Tai, X.C., Chan, T.: Image segmentation using Euler's elastica as the regularization. Journal of Scientific Computing **57**(2), 414–438 (2013)