Iranian Journal of Numerical Analysis and Optimization
Vol. 10, No. 2, (2020), pp 87–104
DOI:10.22067/ijnao.v10i2.84900

Research Article

A four directions variational method for solving image processing problems†

A.R. Hosseini* and E.E. Esfahani

Abstract

In this paper, based on a discrete total variation model, a modified discretization of total variation (TV) is introduced for image processing problems. Two optimization problems corresponding to compressed sensing magnetic resonance imaging (MRI) data reconstruction problem and image denoising are proposed. In the proposed method, instead of applying isotropic TV whose gradient field is a two directions vector, a four directions discretization with some modification is applied for the inverse problems. A dual formulation for the proposed TV is explained and an efficient primal-dual algorithm is employed to solve the problem. Some important image test problems in MRI and image denoising problems are considered in the numerical experiments. We compare our model with the state of the art methods.

AMS(2010): Primary 62H35; Secondary 68U10, 94A08.

Keywords: Total variation; Magnetic resonance imaging; Primal-dual optimization method; Regularization; Image denoising.

1 Introduction

There are different methods to solve image processing problems such as spacial filtering [11, 6], methods based on partial differential equations [21, 26],

*Corresponding author
Received 1 January 2020; revised 1 March 2020; accepted 31 May 2020

Alireza Hosseini
School of Mathematics, Statistics and Computer Science, College of Science, University of Tehran, Iran. e-mail: hosseini.alireza@ut.ac.ir

Erfan Ebrahim Esfahani
School of Mathematics, Statistics and Computer Science, College of Science, University of Tehran, Iran. e-mail: ebrahim.esfahani@ut.ac.ir

† This article was suggested by the Scientific Committee of the Third Iranian Seminar on Control and Optimization for publication in IJNAO, which was accepted after independent review.

variational methods [8, 2, 10, 1, 5], and machine learning approach as well as the methods based on deep learning [17, 25]. Variational methods propose some different functionals as the tools for calculating accidental variations in a signal or function. The pioneer work in this area was proposed by Rudin, Osher, and Fatemi [24]. Recently, Condat [10] proposed an isotropic discrete total variation with high accuracy. Moreover, the total generalized variation as a generalization of total variation functional has been proposed in [5]. Consider the following optimization problem to solve a continuous version of mathematical image problems:

$$\min_s \frac{1}{2}\int_\Omega |G(s(t)) - g(t)|^2 dt + \lambda J(s),$$

where G is an appropriate function (for more details see [7]) and

$$J(s) = \sup\left\{-\int_\Omega s.div\phi dt : \phi \in C_c^1(\Omega, \mathbb{R}^N), |\phi(t)| \leq 1, \forall t \in \Omega\right\}. \quad (1)$$

Also J is the dual definition of total variation (TV) of the $L_c^1(\Omega)$ (locally Lebesgue integrable) function s. If $G = I$, then problem (1) is named denoising problem (for image g). A function s is said to have bounded variation whenever $J(s) < \infty$. The space $BV(\Omega)$ of functions with bounded variation is the set of functions $s \in L^1(\Omega)$ such that $J(s) < \infty$, endowed with the norm $\|s\|_{BV(\Omega)} = \|s\|_{L^1(\Omega)} + J(s)$. Obviously, for the smooth function $s \in C^1(\Omega)$ (or $s \in W^{1,1}(\Omega)$),

$$J(s) = \int_\Omega |\nabla s| dt$$

the minimization of $J(s)$ is equivalent to minimization of the sum of abrupt variations (the majority of the derivative) over its dimension. To solve real digital imaging problems, discretization of (1) is inevitable. There are different forms of discretization for TV functional in the papers. As the review of such literatures is long some, we refer the readers to [10, 13] and references therein.

Total generalized variation (TGV) is a cleverly defined functional for calculating a combinational value of function variation by means of derivatives up to a given order. TGV has some superiorities in comparison with the first-order TV, such as its ability to distinguish staircase artifacts, and as a result, more cleaner image can be reconstructed. Usually in applications, the second-order TGV is sufficient, because the higher-order will face to the complexity of calculation. The second-order TGV is defined by the following optimization problem (see [5]):

$$\mathrm{TGV}_\alpha^2(u) = \sup\left\{\int_\Omega u.\mathrm{div}^2 v dx : v \in C_c^2(\Omega, \mathrm{Sym}^2(\mathbb{R}^d)), \|\mathrm{div}^l v\|_\infty \leq \alpha_l, l = 0, 1\right\}.$$

In [5], a discretization of the TGV was proposed. Based on discretization of TV in [13], in this paper, a new discretization of TV (1) is given; operators are changed and linear operator $\mathcal{O}_+$ is inserted to ensure the projection on the

vertices. Then the magnetic resonance imaging (MRI) data reconstruction and image denoising problems will be formulated based on the new proposed discrete TV and will be compared with some state of the art models. In the proposed model, instead of using two directions gradient field in isotropic TV, a four directions discretization model is proposed. Actually, we can take more information from structure of the images, and consequently a new function for measuring the majority of the noise and oscillation is proposed. Moreover, after some modification based on the Condat's model [10], a discretization of the total variation with better isotropy property will be defined. Numerical results show that the proposed method has better performance in comparison with isotropic TV, TGV, and some other famous models in terms of removing noise and compressed sensing MRI problems.

The paper is arranged as the following; in Section 2, the Chambolle–Pock algorithm is explained for solving optimization problems in image processing. In Section 3, a new discrete TV model is proposed. In Section 4, the problem of reconstruction of undersampled MRI data as well as the problem of image denoising are formulated. Some important test problems in medical MRI and denoising are solved numerically and compared with some state of the art, in Section 4. Finally some conclusions are given in Section 5.

2 A numerical algorithm for solving convex image problems

2.1 Algorithm description

Consider a finite-dimensional vector space X with an inner product $\langle\cdot,\cdot\rangle_X$ and Euclidean norm $||\cdot||_2 = \sqrt{\langle\cdot,\cdot\rangle_X}$. $C \subset X$ we define the *indicator* function of C as

$$\delta_C(x) = \begin{cases} 0, & x \in C, \\ \infty, & x \notin C. \end{cases}$$

For any function $F : X \to [-\infty, +\infty]$, the *convex conjugate* at $z \in X$ is defined by

$$F^*(z) = \max_x \langle x, z\rangle - F(x). \tag{2}$$

Moreover, for any $\theta > 0$ and $\bar{x} \in X$, the *proximal* (or *proximity*) mapping of F at $\bar{x}$ is defined by

$$\text{prox}_{\theta F}(\bar{x}) = \arg\min_x \frac{||x - \bar{x}||^2}{2\theta} + F(x).$$

Now, suppose that Y is another real vector space with inner product $\langle\cdot,\cdot\rangle_Y$ and the corresponding induced norm $||\cdot||_2 = \sqrt{\langle\cdot,\cdot\rangle_Y}$. Let $K : X \to Y$ be a

bounded linear operator with the operator norm

$$||K|| := \max\{||Kx||_2 : x \in X \text{ and } ||x||_2 \leq 1\} < \infty.$$

The adjoint of K is defined as an operator K^* for which the equality $\langle Kx, y\rangle_Y = \langle x, K^*y\rangle_X$ holds for all x and y. The general problem, we tackle, is

$$\min_x F(Kx) + G(x), \tag{3}$$

where $F : X \to [0, \infty)$ and $G : Y \to [0, \infty)$ are proper, closed, and convex. By virtue of (2), the convex minimization problem (3) can be reformulated as the saddle point convex-concave problem

$$\min_x \max_y \langle Kx, y\rangle_Y + G(x) - F^*(y), \tag{4}$$

which is also known as the primal-dual formulation of (3). Hereafter we assume that a solution to this problem always exists. The method we employ to solve (4), which lays the groundwork of what we shall propose in this paper, is the celebrated Chambolle–Pock algorithm, recalled below: The algorithm

Data: K, K^*, F^*, G
Result: x^k
initialization: Choose parameters $\sigma > 0$, $\tau > 0$, and initial estimates $(x^0, y^0) \in X \times Y$.
while *convergence criterion not met, for k=0,1,...,* **do**
  1-$y^{k+1} = \text{prox}_{\sigma F^*}(y^k + \sigma K\tilde{x}^k)$,
  2- $x^{k+1} = \text{prox}_{\tau G}(x^k - \tau K^* y^{k+1})$,
  3-$\tilde{x}^{k+1} = 2x^{k+1} - x^k$.
end

Algorithm 1: The Chambolle–Pock algorithm for solving problem (4)

is guaranteed to converge provided that $\sigma\tau < \frac{1}{||K||^2}$. In practical situations, where computing the exact value of $||K||$ is not easy, finding an upper bound $B \geq ||K||^2$ and setting $\sigma = \tau = \frac{1}{\sqrt{B}}$ are sufficient for convergence. We chose this algorithm over other well-known first-order methods such as ADMM or FISTA because of its superior convergence rate [9], and the fact that it requires fewer parameters to be tuned.

2.2 Convergence analysis

Here, we express the convergence theorem of the Chambolle–Pock algorithm to solve problem (4). The convergence analysis was argued in [9]. The following theorem guarantees the convergence of Chambolle–Pock algorithm for

solving problem (4).

Theorem 1. [9] Assume X and Y are two finite-dimensional Hilbert spaces and that $\tau, \sigma > 0$. Moreover assume $\sigma\tau\|K\|^2 < 1$. Then there exists a solution $(\bar{x}, \bar{y}) \in X \times Y$ of problem (4) such that the sequence (x^n, y^n) in Algorithm 1 converges to $(\bar{x}, \bar{y})$.

Note that, in [9], it was shown that the rate of convergence is $\mathcal{O}(\frac{1}{n})$. Moreover, under some uniformly convexity assumptions, the rate of convergence can be improved to $\mathcal{O}(\frac{1}{e^N})$.

3 A Discretization for TV

Assume that $x \in \mathbb{R}^{N_1 \times N_2}$ is a gray scaled two-dimensional image and that $A = \{(n_1, n_2), n_1 = 1, \ldots, N_1, n_2 = 1, \ldots, N_2\}$ is a set of grids. Consider the following new discrete TV model, which is based on four directional gradient with imposed some operational constraints:

$$\begin{aligned}\mathrm{TV}_{dis}(x) = \max_{u \in (\mathbb{R}^4)^{N_1 \times N_2}} \{ < \mathbb{D}x, u > :\ & |\mathcal{O}_{\updownarrow} u(n_1, n_2)| \le 1,\ |\mathcal{O}_{\leftrightarrow} u(n_1, n_2)| \le 1, \\ & |\mathcal{O}_{\bullet} u(n_1, n_2)| \le 1,\ |\mathcal{O}_{+} u(n_1, n_2)| \le 1, \forall (n_1, n_2) \in X\},\end{aligned}$$

where

$$\begin{aligned}&\mathbb{D} = (D_1, D_2, D_3, D_4), \\ &D_1 x(n_1, n_2) = x(n_1 + 1, n_2) - x(n_1, n_2), \\ &D_2 x(n_1, n_2) = x(n_1, n_2 + 1) - x(n_1, n_2), \\ &D_3 x(n_1, n_2) = x(n_1 + 1, n_2 + 1) - x(n_1, n_2), \\ &D_4 x(n_1, n_2) = x(n_1 - 1, n_2 + 1) - x(n_1, n_2),\end{aligned}$$

$$|\mathcal{O}_{\star} u(n_1, n_2)| = \sqrt{\sum_{j=1}^{4} [(\mathcal{O}_{\star} u)_j (n_1, n_2)]^2}, \quad (n_1, n_2) \in X, \star = \updownarrow, \leftrightarrow, \bullet, +,$$

and

$$\begin{aligned}(\mathcal{O}_{\updownarrow} u)_1(n_1, n_2) &= u_1(n_1, n_2), \\ (\mathcal{O}_{\updownarrow} u)_2(n_1, n_2) &= \tfrac{1}{4}(u_2(n_1, n_2) + u_2(n_1, n_2 - 1) \\ &\quad + u_2(n_1 + 1, n_2) + u_2(n_1 + 1, n_2 - 1)), \\ (\mathcal{O}_{\updownarrow} u)_3(n_1, n_2) &= \tfrac{1}{2}(u_3(n_1, n_2) + u_3(n_1, n_2 - 1)), \\ (\mathcal{O}_{\updownarrow} u)_4(n_1, n_2) &= \tfrac{1}{4}(u_4(n_1, n_2) + u_4(n_1, n_2 - 1) \\ &\quad + u_4(n_1 + 1, n_2) + u_4(n_1 + 1, n_2 - 1),\end{aligned}$$

$$
\begin{aligned}
(\mathcal{O}_{\leftrightarrow}u)_1(n_1,n_2) &= \tfrac{1}{4}(u_1(n_1,n_2)+u_1(n_1-1,n_2)\\
&\quad +u_1(n_1,n_2+1)+u_1(n_1-1,n_2+1)),\\
(\mathcal{O}_{\leftrightarrow}u)_2(n_1,n_2) &= u_2(n_1,n_2),\\
(\mathcal{O}_{\leftrightarrow}u)_3(n_1,n_2) &= \tfrac{1}{2}(u_3(n_1,n_2)+u_3(n_1-1,n_2)),\\
(\mathcal{O}_{\leftrightarrow}u)_4(n_1,n_2) &= \tfrac{1}{2}(u_4(n_1,n_2)+u_4(n_1+1,n_2)),
\end{aligned}
$$

$$
\begin{aligned}
(\mathcal{O}_{+}u)_1(n_1,n_2) &= \tfrac{1}{2}(u_1(n_1,n_2)+u_1(n_1,n_2+1)),\\
(\mathcal{O}_{+}u)_2(n_1,n_2) &= \tfrac{1}{2}(u_2(n_1,n_2)+u_2(n_1+1,n_2)),\\
(\mathcal{O}_{+}u)_3(n_1,n_2) &= u_3(n_1,n_2),\\
(\mathcal{O}_{+}u)_4(n_1,n_2) &= u_4(n_1,n_2),
\end{aligned}
$$

$$
\begin{aligned}
(\mathcal{O}_{\bullet}u)_1(n_1,n_2) &= \tfrac{1}{2}(u_1(n_1,n_2)+u_1(n_1-1,n_2)),\\
(\mathcal{O}_{\bullet}u)_2(n_1,n_2) &= \tfrac{1}{2}(u_2(n_1,n_2)+u_2(n_1,n_2-1)),\\
(\mathcal{O}_{\bullet}u)_3(n_1,n_2) &= \tfrac{1}{4}(u_3(n_1,n_2)+u_3(n_1,n_2-1)\\
&\quad +u_3(n_1-1,n_2)+u_3(n_1-1,n_2-1)),\\
(\mathcal{O}_{\bullet}u)_4(n_1,n_2) &= \tfrac{1}{4}(u_4(n_1,n_2)+u_4(n_1+1,n_2)\\
&\quad +u_4(n_1,n_2-1)+u_4(n_1+1,n_2-1)).
\end{aligned}
$$

Theorem 2. Assume that F is convex and that $\lambda \geq 0$. Then the problem

$$\min_x F(x) + \lambda \mathrm{TV}_{dis}(x),$$

is equivalent to the following optimization problem:

$$
\begin{aligned}
&\min F(x) + \lambda\{|v_{\updownarrow}| + |v_{\leftrightarrow}| + |v_{\bullet}| + |v_{+}|\},\\
&\text{subject to}\\
&\mathcal{O}^*_{\updownarrow}(v_{\updownarrow})(n_1,n_2) + \mathcal{O}^*_{\leftrightarrow}(v_{\leftrightarrow})(n_1,n_2) + \mathcal{O}^*_{\bullet}(v_{\bullet})(n_1,n_2)\\
&\qquad\qquad +\mathcal{O}^*_{+}(v_{+})(n_1,n_2) = \mathbb{D}x(n_1,n_2), \quad (n_1,n_2) \in X,
\end{aligned}
$$

where

$$
\mathcal{O}^*_{\updownarrow}(v_{\updownarrow})(n_1,n_2) = \begin{pmatrix}
v^1_{\updownarrow}(n_1,n_2)\\
\\
\tfrac{1}{4}(v^2_{\updownarrow}(n_1,n_2)+v^2_{\updownarrow}(n_1,n_2+1)+v^2_{\updownarrow}(n_1-1,n_2)\\
\qquad +v^2_{\updownarrow}(n_1-1,n_2+1))\\
\\
\tfrac{1}{2}(v^3_{\updownarrow}(n_1,n_2)+v^3_{\updownarrow}(n_1,n_2+1))\\
\\
\tfrac{1}{4}(v^4_{\updownarrow}(n_1,n_2)+v^4_{\updownarrow}(n_1,n_2+1)+v^4_{\updownarrow}(n_1-1,n_2)\\
\qquad +v^4_{\updownarrow}(n_1-1,n_2+1))
\end{pmatrix},
$$

$$\mathcal{O}^*_{\leftrightarrow}(v_{\leftrightarrow})(n_1,n_2)=\begin{pmatrix}\frac{1}{4}(v^1_{\leftrightarrow}(n_1,n_2)+v^1_{\leftrightarrow}(n_1+1,n_2)+v^1_{\leftrightarrow}(n_1,n_2-1)\\ +v^1_{\leftrightarrow}(n_1+1,n_2-1))\\ \\ v^2_{\leftrightarrow}(n_1,n_2)\\ \\ \frac{1}{2}(v^3_{\leftrightarrow}(n_1,n_2)+v^3_{\leftrightarrow}(n_1+1,n_2))\\ \\ \frac{1}{2}(v^4_{\leftrightarrow}(n_1,n_2)+v^4_{\leftrightarrow}(n_1-1,n_2))\end{pmatrix},$$

$$\mathcal{O}^*_{\bullet}(v_{\leftrightarrow})(n_1,n_2)=\begin{pmatrix}\frac{1}{2}(v^1_{\bullet}(n_1,n_2)+v^1_{\bullet}(n_1+1,n_2))\\ \\ \frac{1}{2}(v^2_{\bullet}(n_1,n_2)+v^2_{\bullet}(n_1,n_2+1))\\ \\ \frac{1}{4}(v^3_{\bullet}(n_1,n_2)+v^3_{\bullet}(n_1,n_2+1)+v^3_{\bullet}(n_1+1,n_2)\\ +v^3_{\bullet}(n_1+1,n_2+1))\\ \\ \frac{1}{4}(v^4_{\bullet}(n_1,n_2)+v^4_{\bullet}(n_1-1,n_2)+v^4_{\bullet}(n_1,n_2+1)\\ +v^4_{\bullet}(n_1-1,n_2+1))\end{pmatrix}$$

$$\mathcal{O}^*_{+}(v_{+})(n_1,n_2)=\begin{pmatrix}\frac{1}{2}(v^1_{+}(n_1,n_2)+v^1_{+}(n_1,n_2-1))\\ \\ \frac{1}{2}(v^2_{+}(n_1,n_2)+v^1_{+}(n_1-1,n_2))\\ \\ v^3_{+}(n_1,n_2)\\ \\ v^4_{+}(n_1,n_2)\end{pmatrix},$$

and

$$|v_\star|=\sqrt{\sum_{n_1,n_2}|v^j_\star|^2},\star=\updownarrow,\leftrightarrow,\bullet,+.$$

Proof. It is a direct conclusion of Fenchel–Rockefeller, primal-dual theorem (see [10], for instance). □

4 Formulation of some image processing problems

4.1 Compressing MRI data

Consider the following problem for reconstructing MR images from undersampled k-space data:

$$\min_{u\in U}\frac{1}{2\lambda}||\mathcal{F}_m u-b||_2^2+\beta||\Psi u||_1+\mathrm{TV}_{dis}(u), \tag{5}$$

where $u \in U = \mathbb{R}^{n\times n}$ is the MR image, which is defined in spatial domain and $\mathcal{F}_m$ is the undersampled Fourier operator, which converts signals in spatial domain to its corresponding frequency representation and undersample it. In other words, it can be defined as $\mathcal{F}_m u := M \odot \mathcal{F}(u)$, where $\mathcal{F}$ is the two-dimensional fast Fourier transform, $M_{n\times n}$ is the k-space sampling mask with ones at m sampled frequencies and zeros at undersampled locations, and $\odot$ is the component-wise multiplication. We further assume that the sampling rate $\frac{m}{n^2}$ is considerably small (highly undersampled k-space) and that $b \in \mathbb{C}^{n\times n}$ is the partially scanned k-space with only m sampled frequencies and $n^2 - m$ zeros. The so-called zero-filled solution is defined as $u_{\text{zf}} := \mathcal{F}^{-1}(b)$. Furthermore, Ψ is a discrete wavelet transform, and the constants β, $\lambda \in \mathbb{R}$ are regularization parameters.
Referring to Theorem 2, we can get the following equivalent formulation:

$$\begin{aligned} &\min_{u\in U} \frac{1}{2\lambda}||\mathcal{F}_m u - b||_2^2 + \beta||\Psi u||_1 + |v_{\updownarrow}^*|_1 + |v_{\leftrightarrow}^*|_1 + |v_{\bullet}^*|_1 + |v_{+}^*|_1, \\ &\text{s.t. } \mathcal{O}_{\updownarrow}^*(v_{\updownarrow})(n_1,n_2) + \mathcal{O}_{\leftrightarrow}^*(v_{\leftrightarrow})(n_1,n_2) + \mathcal{O}_{\bullet}^*(v_{\bullet})(n_1,n_2) \\ &\qquad\qquad + \mathcal{O}_{+}^*(v_{+})(n_1,n_2) = \mathbb{D}u(n_1,n_2), \quad (n_1,n_2) \in A. \\ &A = \{(n_1,n_2) : n_1,n_2 = 1,\ldots,n\}. \end{aligned} \tag{6}$$

To solve this problem with Algorithm 1, we need to reformulate problem (6) to the form of (3). Now we define the following linear operator K:

$$Kx = \begin{pmatrix} \Psi & 0 & 0 & 0 & 0 \\ \mathcal{F}_m & 0 & 0 & 0 & 0 \\ -\mathbb{D} & \mathcal{O}_{\updownarrow}^* & \mathcal{O}_{\leftrightarrow}^* & \mathcal{O}_{\bullet}^* & \mathcal{O}_{+}^* \end{pmatrix} \begin{pmatrix} u \in \mathbb{R}^{N_1\times N_2} \\ v_{\updownarrow} \in (\mathbb{R}^4)^{N_1\times N_2} \\ v_{\leftrightarrow} \in (\mathbb{R}^4)^{N_1\times N_2} \\ v_{\bullet} \in (\mathbb{R}^4)^{N_1\times N_2} \\ v_{+} \in (\mathbb{R}^4)^{N_1\times N_2} \end{pmatrix} = \begin{pmatrix} \bar{k} \in (\mathbb{R})^{N_1\times N_2} \\ \bar{h} \in (\mathbb{R})^{N_1\times N_2} \\ \bar{v} \in (\mathbb{R}^4)^{N_1\times N_2} \end{pmatrix} = Y.$$

Assume

$$F(Y) = \frac{1}{2\lambda}\|Y_2 - b\|_2^2 + \beta\|Y_1\|_1 + \delta_{\{0\}}(Y_3),$$
$$G(X) = |v_{\updownarrow}^*|_1 + |v_{\leftrightarrow}^*|_1 + |v_{\bullet}^*|_1 + |v_{+}^*|_1.$$

From arguments about calculating adjoint of a given operator, we get

$$F^*(y) = \frac{\lambda}{2}\|y_2\|_2^2 + < y_2, b > + \delta_{\|\cdot\|_\infty \le \beta}(y_1) + < y_3, \mathbb{D}u > .$$

4.1.1 The *y*-subproblems

According to Algorithm 1, for $\{Y\} = (\bar{k}, \bar{h}, \bar{w})^\top$, we obtain

Result: u^k
Initialization: Set parameters as $\sigma = \tau = \frac{1}{\sqrt{19}}$ and initial estimates as $u^0 = u_{\text{zf}}$, $\tilde{u}^0 = 0$, $v_{\updownarrow}^0 = \tilde{v}_{\updownarrow}^0 = 0$, $v_{\leftrightarrow}^0 = \tilde{v}_{\leftrightarrow}^0 = 0$, $v_{\bullet}^0 = \tilde{v}_{\bullet}^0 = 0$, $v_{+}^0 = \tilde{v}_{+}^0 = 0$, $\bar{h}^0 = 0$, $\bar{k}^0 = 0$ and $\bar{w}^0 = 0$.
while *convergence criterion not met, for k=0,1,...* **do**
1-$\bar{h}^{k+1} = \text{prox}_{\sigma(\frac{\lambda}{2}||\cdot||_2^2 + .^t b)}\big(\bar{h}^k + \sigma(\mathcal{F}_m(\tilde{u}^k))\big), \text{prox}_{\sigma(\frac{\lambda}{2}||\cdot||_2^2 + .^t b)}(\alpha) = \frac{\alpha - \sigma b}{1+\sigma\lambda}$
2-$\bar{k}^{k+1} = \text{proj}_{||\cdot||_\infty \leq \beta}\big(\bar{k}^k + \sigma(\Psi\tilde{u}^k)\big), \Big(\text{proj}_{||\cdot|| \leq \beta}(s)\Big)(i,j) =$
$\min\{\max\{s(i,j), -\beta\}, \beta\}, \quad \forall i,j \in \{1,2,\ldots,n\},$
3-$\bar{w}^{k+1} = \bar{w}^k + \sigma(\mathcal{O}_{\updownarrow}^* \tilde{v}_{\updownarrow}^k + \mathcal{O}_{\leftrightarrow}^* \tilde{v}_{\leftrightarrow}^k + \mathcal{O}_{\bullet}^* \tilde{v}_{\bullet}^k + \tilde{v}_{+}^k - \mathbb{D}\tilde{u}^k)$
4-$u^{k+1} = u^k - \tau(\mathcal{F}_m^* \bar{h}^{k+1} + \Psi^* \bar{k}^{k+1} - \mathbb{D}^* \bar{w}^{k+1})$,
Use formula $\text{Prox}_{\theta\alpha_0|\cdot|_1}(\alpha)(i,j) = (1 - \frac{\theta\alpha_0}{\max\{\|\alpha(i,j)\|_2, \theta\alpha_0\}})\alpha(i,j)$ for steps 5-8 below:
5-$v_{\updownarrow}^{k+1} = \text{Prox}_{\tau|\cdot|_1}(v_{\updownarrow}^k - \tau(\mathcal{O}_{\updownarrow}\bar{w}^{k+1}))$,
6-$v_{\leftrightarrow}^{k+1} = \text{Prox}_{\tau|\cdot|_1}(v_{\leftrightarrow}^k - \tau(\mathcal{O}_{\leftrightarrow}\bar{w}^{k+1}))$,
7-$v_{\bullet}^{k+1} = \text{Prox}_{\tau|\cdot|_1}(v_{\bullet}^k - \tau(\mathcal{O}_{\bullet}\bar{w}^{k+1}))$,
8-$v_{+}^{k+1} = \text{Prox}_{\tau|\cdot|_1}(v_{+}^k - \tau(\mathcal{O}_{+}\bar{w}^{k+1}))$,
9-$\tilde{u}^{k+1} = 2u^{k+1} - u^k, \tilde{v}_{\updownarrow}^{k+1} = 2v_{\updownarrow}^{k+1} - v_{\updownarrow}^k, \tilde{v}_{\leftrightarrow}^{k+1} = 2v_{\leftrightarrow}^{k+1} - v_{\leftrightarrow}^k, \tilde{v}_{\bullet}^{k+1} =$
$2v_{\bullet}^{k+1} - v_{\bullet}^k, \tilde{v}_{+}^{k+1} = 2v_{+}^{k+1} - v_{+}^k.$
end

Algorithm 2: The algorithm for undersampled MRI reconstruction

$$\begin{pmatrix} \bar{k}^{k+1} \\ \bar{h}^{k+1} \\ \bar{w}^{k+1} \end{pmatrix} = \text{prox}_{\sigma F^*}\left(\begin{pmatrix} \bar{k}^k \\ \bar{h}^k \\ \bar{w}^k \end{pmatrix} + \sigma \begin{pmatrix} \Psi & 0 & 0 & 0 & 0 \\ \mathcal{F}_m & 0 & 0 & 0 & 0 \\ -\mathbb{D} & \mathcal{O}_{\updownarrow}^* & \mathcal{O}_{\leftrightarrow}^* & \mathcal{O}_{\bullet}^* & \mathcal{O}_{+}^* \end{pmatrix} \begin{pmatrix} u \in \mathbb{R}^{N_1 \times N_2} \\ v_{\updownarrow} \in (\mathbb{R}^4)^{N_1 \times N_2} \\ v_{\leftrightarrow} \in (\mathbb{R}^4)^{N_1 \times N_2} \\ v_{\bullet} \in (\mathbb{R}^4)^{N_1 \times N_2} \\ v_{+} \in (\mathbb{R}^4)^{N_1 \times N_2} \end{pmatrix}\right)$$
$$= \text{prox}_{\sigma F^*}\begin{pmatrix} \bar{k}^k + \sigma(\Psi(u^k)) \\ \bar{h}^k + \sigma(\mathcal{F}_m \tilde{u}^k) \\ \bar{w}^k + \sigma(\mathcal{O}_{\updownarrow}^* v_{\updownarrow}^k + \mathcal{O}_{\leftrightarrow}^* v_{\leftrightarrow}^k + \mathcal{O}_{\bullet}^* v_{\bullet}^k + \mathcal{O}_{+}^* v_{+}^k) \end{pmatrix}.$$

$$\begin{cases} \bar{h}^{k+1} = \text{prox}_{\sigma(\frac{\lambda}{2}||\cdot||_2^2 + .^t b)}\big(\bar{h}^k + \sigma(\mathcal{F}_m(u^k))\big), \text{prox}_{\sigma(\frac{\lambda}{2}||\cdot||_2^2 + .^t b)}(\alpha) = \frac{\alpha - \sigma b}{1+\sigma\lambda}, \\ \bar{k}^{k+1} = \text{proj}_{||\cdot||_\infty \leq \beta}\big(\bar{k}^k + \sigma(\Psi\tilde{u}^k)\big), \Big(\text{proj}_{||\cdot|| \leq \beta}(s)\Big)(i,j) \\ \qquad = \min\{\max\{s(i,j), -\beta\}, \beta\}, \qquad \forall i,j \in \{1,2,\ldots,n\}, \\ \bar{w}^{k+1} =< \bar{w}^k + \sigma(\mathcal{O}_{\updownarrow}^* v_{\updownarrow}^k + \mathcal{O}_{\leftrightarrow}^* v_{\leftrightarrow}^k + \mathcal{O}_{\bullet}^* v_{\bullet}^k + \mathcal{O}_{+}^* v_{+}^k - \mathbb{D}u). \end{cases}$$

4.1.2 The x-subproblems

For $x = (u, v_{\updownarrow}, v_{\leftrightarrow}, v_{\bullet}, v_{+})^\top$, according to Algorithm 1, we get,

$$\begin{pmatrix} u^{k+1} \\ v_{\updownarrow}^{k+1} \\ v_{\leftrightarrow}^{k+1} \\ v_{\bullet}^{k+1} \\ v_{+}^{k+1} \end{pmatrix} = \text{prox}_{\tau G} \left(\begin{pmatrix} u^{k} \\ v_{\updownarrow}^{k} \\ v_{\leftrightarrow}^{k} \\ v_{\bullet}^{k} \\ v_{+}^{k} \end{pmatrix} - \tau \begin{pmatrix} \Psi^* & \mathcal{F}_m^* & -\mathbb{D}^* \\ 0 & 0 & \mathcal{O}_{\updownarrow} \\ 0 & 0 & \mathcal{O}_{\leftrightarrow} \\ 0 & 0 & \mathcal{O}_{\bullet} \\ 0 & 0 & \mathcal{O}_{+} \end{pmatrix} \begin{pmatrix} \bar{k}^{k+1} \\ \bar{h}^{k+1} \\ \bar{w}^{k+1} \end{pmatrix} \right)$$

$$= \text{prox}_{\tau G} \begin{pmatrix} u^k - \tau(\Psi^* \bar{k}^{k+1} + \mathcal{F}_m^* \bar{h}^{k+1} - \mathbb{D}^* \bar{w}^{k+1}) \\ v_{\updownarrow}^k - \tau(\mathcal{O}_{\updownarrow} \bar{w}^{k+1}) \\ w_{\leftrightarrow}^k - \tau(\mathcal{O}_{\leftrightarrow} \bar{w}^{k+1}) \\ w_{\bullet}^k - \tau(\mathcal{O}_{\bullet} \bar{w}^{k+1}) \\ w_{+}^k - \tau(\mathcal{O}_{+} \bar{w}^{k+1}) \end{pmatrix},$$

where

$$\begin{aligned} \mathbb{D}^* u(n_1, n_2) = &[u_1(n_1 - 1, n_2) - u_1(n_1, n_2)] + [u_2(n_1 - 1, n_2 - 1) - u_2(n_1, n_2)] \\ &+[u_3(n_1 - 1, n_2 - 1) - u_3(n_1, n_2)] \\ &+[u_4(n_1 + 1, n_2 - 1) - u_4(n_1, n_2)], \quad (n_1, n_2) \in A. \end{aligned}$$

Based on the above argument, we propose Algorithm 2, to solve undersampled MRI reconstruction problem.

4.2 Image denoising formulation

Let $z \in \mathbb{R}^{N_1 \times N_2}$ be a gray scaled two-dimensional noisy image, and consider the following optimization problem:

$$\min_u \frac{1}{2} \|z - u\|_2^2 + \lambda \text{TV}_{dis}(u).$$

This problem is named denoising problem and can be rewritten with the following problem:

$$\begin{aligned} &\min_{u \in U} \frac{1}{2} ||z - u||_2^2 + |v_{\updownarrow}^*|_1 + |v_{\leftrightarrow}^*|_1 + |v_{\bullet}^*|_1 + |v_{+}^*|_1 \\ &\text{s.t. } \mathcal{O}_{\updownarrow}^*(v_{\updownarrow})(n_1, n_2) + \mathcal{O}_{\leftrightarrow}^*(v_{\leftrightarrow})(n_1, n_2) + \mathcal{O}_{\bullet}^*(v_{\bullet})(n_1, n_2) + \mathcal{O}_{+}^*(v_{+})(n_1, n_2) \\ &\qquad = \mathbb{D}u(n_1, n_2), \quad (n_1, n_2) \in X. \end{aligned}$$

To solve this problem numerically, we need to determine functions F and G as well as the linear operator K in (3). Assume $x = (u, v_{\updownarrow}^*, v_{\leftrightarrow}^*, v_{\bullet}^*, v_{+}^*)^\top$. Set

$$K = \begin{pmatrix} I & 0 & 0 & 0 & 0 \\ -\mathbb{D} & \mathcal{O}_{\updownarrow}^* & \mathcal{O}_{\leftrightarrow}^* & \mathcal{O}_{\bullet}^* & \mathcal{O}_{+}^* \end{pmatrix}, \qquad FY = \frac{1}{2} \|Y_1 - z\|_2^2 + \delta_{\{0\}}(Y_2 - \mathbb{D}z),$$

$$Gx = |v_{\updownarrow}^*|_1 + |v_{\leftrightarrow}^*|_1 + |v_{\bullet}^*|_1 + |v_{+}^*|_1.$$

Consequently

$$K^* = \begin{pmatrix} I & -\mathbb{D}^* \\ 0 & \mathcal{O}_{\updownarrow} \\ 0 & \mathcal{O}_{\leftrightarrow} \\ 0 & \mathcal{O}_{\bullet} \\ 0 & \mathcal{O}_{+} \end{pmatrix}, \qquad F^* y = \|y_1\|_2^2 + < z, y_1 > .$$

Analogous to the arguments about Algorithm 2, we can obtain the similar procedure (Algorithm 3), to solve the image denoising problem.

Result: u^k
Initialization: Set parameters as $\sigma = \tau = \frac{1}{\sqrt{19}}$ and initial estimates as $u^0 = z$, $\tilde{u}^0 = 0$, $v_{\updownarrow}^0 = \tilde{v}_{\updownarrow}^0 = 0$, $v_{\leftrightarrow}^0 = \tilde{v}_{\leftrightarrow}^0 = 0$, $v_{\bullet}^0 = \tilde{v}_{\bullet}^0 = 0$, $v_{+}^0 = \tilde{v}_{+}^0 = 0$, $\bar{h}^0 = 0$ and $\bar{w}^0 = 0$.
while *convergence criterion not met, for k=0,1,...* **do**
1-$\bar{h}^{k+1} = \text{prox}_{\sigma(\frac{1}{2}||\cdot||_2^2 + \cdot^t z)}\left(\bar{h}^k + \sigma(\tilde{u}^k)\right), \text{prox}_{\sigma(\frac{1}{2}||\cdot||_2^2 + \cdot^t z)}(\alpha) = \frac{\alpha - \sigma z}{1+\sigma}$,
2-$\bar{w}^{k+1} = \bar{w}^k + \sigma(\mathcal{O}_{\updownarrow}^* \tilde{v}_{\updownarrow}^k + \mathcal{O}_{\leftrightarrow}^* \tilde{v}_{\leftrightarrow}^k + \mathcal{O}_{\bullet}^* \tilde{v}_{\bullet}^k + \mathcal{O}_{+}^* \tilde{v}_{+}^k - \mathbb{D}\tilde{u}^k)$,
3-$u^{k+1} = u^k - \tau(\bar{h}^{k+1} - \mathbb{D}^* \bar{w}^{k+1})$,
Use formula $\text{Prox}_{\theta\alpha_0|\cdot|_1}(\alpha)(i,j) = (1 - \frac{\theta\alpha_0}{\max\{\|\alpha(i,j)\|_2, \theta\alpha_0\}})\alpha(i,j)$ for steps 4-7 below:
4-$v_{\updownarrow}^{k+1} = \text{Prox}_{\tau|\cdot|_1}(v_{\updownarrow}^k - \tau(\mathcal{O}_{\updownarrow} \bar{w}^{k+1}))$,
5-$v_{\leftrightarrow}^{k+1} = \text{Prox}_{\tau|\cdot|_1}(v_{\leftrightarrow}^k - \tau(\mathcal{O}_{\leftrightarrow} \bar{w}^{k+1}))$,
6-$v_{\bullet}^{k+1} = \text{Prox}_{\tau|\cdot|_1}(v_{\bullet}^k - \tau(\mathcal{O}_{\bullet} \bar{w}^{k+1}))$,
7-$v_{+}^{k+1} = \text{Prox}_{\tau|\cdot|_1}(v_{+}^k - \tau(\mathcal{O}_{+} \bar{w}^{k+1}))$,
8-$\tilde{u}^{k+1} = 2u^{k+1} - u^k$,
9-$\tilde{v}_{\updownarrow}^{k+1} = 2v_{\updownarrow}^{k+1} - v_{\updownarrow}^k$,
10-$\tilde{v}_{\leftrightarrow}^{k+1} = 2v_{\leftrightarrow}^{k+1} - v_{\leftrightarrow}^k$,
11-$\tilde{v}_{\bullet}^{k+1} = 2v_{\bullet}^{k+1} - v_{\bullet}^k$,
12-$\tilde{v}_{+}^{k+1} = 2v_{+}^{k+1} - v_{+}^k$.
end

Algorithm 3: The algorithm for denoising

5 Numerical results

5.1 Compressed MRI data reconstruction problem

In this section, we run the proposed method on some test images and compare them with other well-known models. We choose a 256×256 in-vivo MR image, which is a T_2-weighted axial brain scan. The image was chosen based on its variety of displayed feature and the fraction of FOV they occupy (about 46 percent of the FOV). Based on these sizes, we chose to sample 10 percent of the k-space of the images. The random sampling masks were chosen as variable-density patterns, with more samples in the low-frequency area (in a neighborhood of the origin) and fewer ones towards the periphery of the k-space. We used the Sparse MRI MATLAB package [18] to generate the

masks. We compared the performance of our method with the baseline zero-filling solution, TV + wavelet model and the TGV model. The TV + wavelet model solves the unconstrained minimization problem

$$\min_{u\in U} \frac{1}{2}||\mathcal{F}_m u - b||_2^2 + \beta||\Psi u||_1 + \alpha||\nabla u||_{1,2}.$$

This model was first introduced in [19] and then was further investigated in [20, 14, 16]. We followed the experiment setup proposed in these papers in our tests. In particular, we set $\alpha = 0.001$ and $\beta = 0.035$. The TGV method has the form

$$\min_{u\in U} \frac{1}{2\lambda}||\mathcal{F}_m u - b||_2^2 + \alpha_1||\nabla u - v||_{1,2} + \alpha_0||\mathcal{E}v||_{1,F}$$

and was proposed in [15]. Again, in order to stay true to the original paper, we chose the parameters as in [15]. In particular, we set $\lambda = 8 \times 10^{-5}$, $\alpha_1 = 1$, and $\alpha_0 = 2$. and the Chambolle–Pock algorithm is applied to them. In the proposed model (5), we also fix $\lambda = 8 \times 10^{-5}$ and $\beta = 0.75$ across the simulations. The 4-tap Daubechies wavelet transform was used the experiment and was implemented through the Sparse MRI package. The algorithm ran for 500 iterations.

We compare the quality of output images of each method based on SNR, SSIM, and HFEN indices. Let u_0 be the ground-truth image. The SNR is defined as

$$\mathrm{SNR}(u) := 20 \log \frac{||u_0||_2}{||u - u_0||_2}$$

and the high SNR is related to low MSE. The SSIM index is defined as

$$\mathrm{SSIM}(u) := \frac{(2\mu_u\mu_{u_0} + c_1)(2\sigma_{uu_0} + c_2)}{(\mu_u^2 + \mu_{u_0}^2 + c_1)(\sigma_u^2 + \sigma_{u_0^2} + c_2)},$$

where μ_u is the average of u, σ_u is the standard deviation of u, σ_{uu_0} is the cross-covariance for u and u_0, and $c_1 = (0.01 \times 255)^2$ and $c_2 = (0.03 \times 255)^2$ are regularizing constants. Another newly-proposed error metric is the HFEN (high frequency error norm), which aims at quantifying the quality of reconstructed edges and fine features [22]. This index is defined as

$$\mathrm{HFEN}(u) := \frac{||\mathrm{LoG}(u) - \mathrm{LoG}(u_0)||_2}{||\mathrm{LoG}(u_0)||_2},$$

where LoG is a rotationally symmetric Laplacian of Gaussian filter with a kernel size of 15×15 pixels and standard deviation of 1.5 pixels. We note that the perfect reconstruction has SNR $= \infty$, that SSIM $= 1$ and that HFEN $= 0$. We should also remark that none of these metrics perfectly quantify the visual perception of reconstructed images; although rarely the case, it might

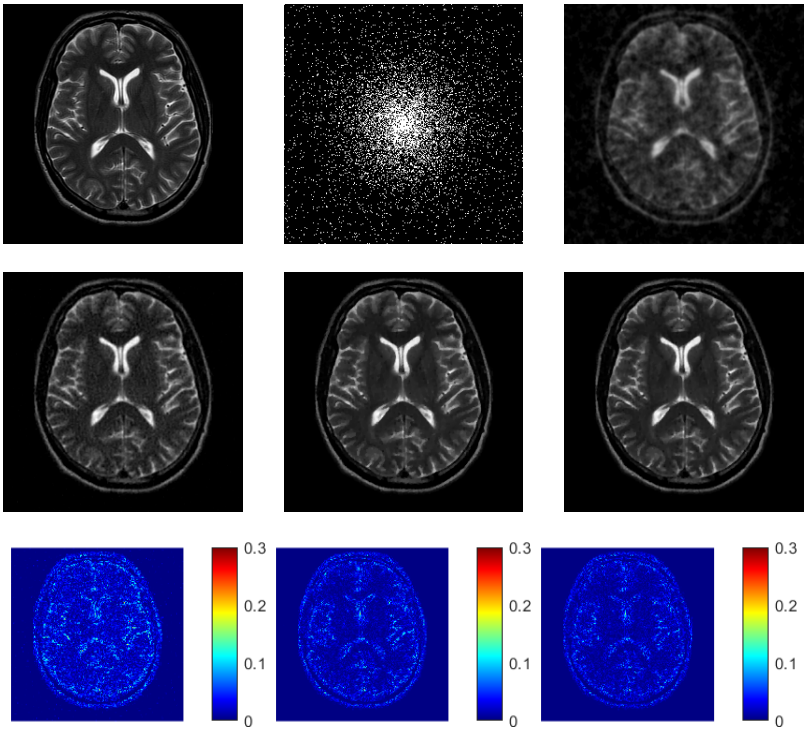

Figure 1: T_2-weighted axial brain scan results. Top row: ground truth (left), sampling mask at rate %10 (middle), zero-filling solution (right). Middle row: reconstruction by TV+W (left), TGV (middle) and the proposed method (right). Bottom row: error maps for TV+W (left), TGV (middle) and the proposed method (right).

just happen that between two reconstructions, one has a higher value in all of these metrics, yet the other is visually more faithful to the ground-truth.

In Figure 1, the reconstruction results for the T_2-weighted brain scan are presented. As expected, the zero-filling solution yields the worst reconstruction with too many incoherent artifacts. The TV+W model removes many of these artifacts but still has low structural similarity to the ground truth, which is also easily observed in the reconstructed image. The TGV model significantly improves upon TV + wavelet model as can be perceived visually and quantitatively (Table 1). The proposed framework moderately improves upon the TGV solution as attested by all three error measures and the error maps (the third row in Figure 1), which shows fewer errors in $[0, 0.1]$ interval.

Table 1: Quantitative comparison between various reconstructions for the T_2-weighted brain scan.

Method	SNR [dB]	SSIM	HFEN
Proposed	**20.71**	**0.9046**	0.1013
TGV	20.07	0.9037	**0.0995**
TV+W	17.58	0.5113	0.2259
ZF	8.83	0.2733	0.7574

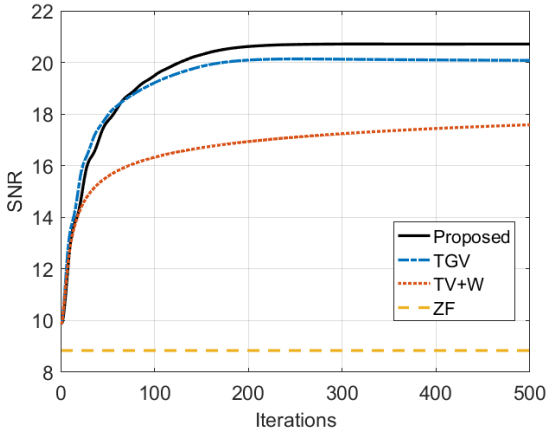

Figure 2: SNR plots for the T_2-weighted brain scan reconstructions.

5.2 Denoising problem

Consider denoising problem (4.2), assume that x is a clean grayscale image and z is a noisy image that is created by adding Gaussian noise with variance 0.1 to the clean image x. In fact, we create an artificial noisy image by means of adding noise to the original image. Here we solve the denoising problem for the `pirate` test image and compare the results of the proposed model, with Condat [10], TGV [5], and Hosseini's model [13]. Table 2 shows the denoised results. It can be seen that the proposed method has better performance in decreasing noise and preserving the details in comparison with the other methods. On the other hand, Figure 3 shows SSIM and PSNR values with respect to the steps. As a result, it can be seen that the proposed model has better values of SSIM and PSNR, at the solutions that are obtained in the convergence.

Table 2: Results of Pirate denoising problem.

Reference	Noisy	Condat
		
TGV	**Hosseini**	**Proposed**
		

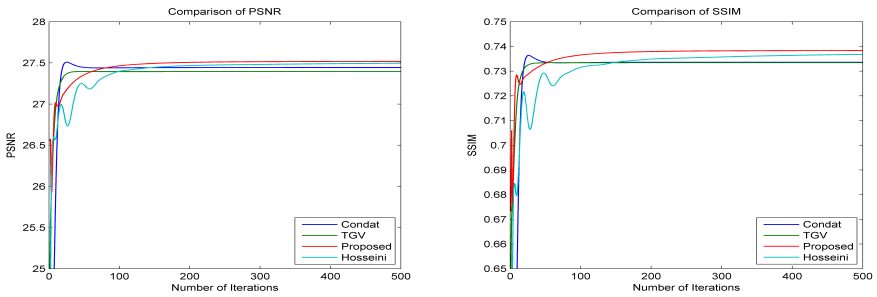

Figure 3: PSNR (left) and SSIM (right) values versus number of iterations for denoising problem.

6 Conclusions

In this paper, a new discretization of total variation functional was proposed. A four directions discrete gradient was used in the structure of the method. Moreover, some linear operators were inserted to the constraints. One of differences of the proposed method with the method in [13], is the new inserted projecting operators to the corners of the pixels. The proposed method

was employed to formulate some image processing problems; reconstructing highly undersampled MR images, and image denoising. The Chambolle–Pock algorithm was applied for the numerical simulations. The wavelet transform was chosen in the formulation of compressed sensing MRI problem and it was compared with zero-filling solution, TGV, and TV+W. In addition, for the image denoising problem, the proposed model was compared with some state of the art methods. In both numerical experiments, the new methods illustrated better quantitative and qualitative results.

References

1. Abergel, R. and Moisan, L. *The Shannon total variation*, J. Math. Imaging Vision 59 (2017), 341–370.
2. Alter, F., Caselles, V. and Chambolle, A. *Evolution of characteristic functions of convex sets in the plane by the minimizing total variation flow*, Interfaces Free Bound. 7 (2005), 29–53 .
3. Beck, A. *First-order methods in optimization,* SIAM, Philadelphia, 2017.
4. Bredies, K. *Recovering piecewise smooth multichannel images by minimization of convex functional with total generalized variation penal,* A. Bruhn et al. eds., Global Optimization Methods, Lecture Notes in Computer Science 8293, (2014), 44–77.
5. Bredies, K., Kunisch, K. and Pock, T. *Total generalized variation*, SIAM J. Imaging Sci. 3 (2010), 492–526.
6. Buades, A., Coll, B. and Morel, J.M. *Image denoising methods. A new nonlocal principle,* SIAM Rev. 52 (2010), 113–147.
7. Chambolle, A., Caselles, V., Novaga, M., Cremers, D. and Pock, T. *An introduction to total variation for image analysis*, Theoretical Foundations and Numerical Methods for Sparse Recovery, De Gruyter, Radon Series Comp. Appl. Math. 9 (2010), 263–340.
8. Chambolle, A., Levine, S.E. and Lucier, B.J. *An upwind finite-difference method for total variation based image smoothing*, SIAM J. Imaging Sci., 4 (2011), 277–299.
9. Chambolle, A. and Pock, T. *A first-order primal-dual algorithm for convex problems with applications to imaging,* J. Math. Imaging Vis. 40 (2010), 120–145.
10. Condat, L. *A primal-dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms*, J. Optim. Theory Appl. 158, 460–479 (2013).

11. Ghazel, M., Freeman, G.H. and Vrscay, E.R. *Fractal image denoising.* IEEE Trans. Image Process. 12 (2003), 1560–1578.

12. Guo, W., Kin, J. and Yin, W. *A new detail-preserving regularization scheme,* SIAM J. Imaging Sci. 7 (2014), 1309–1334.

13. Hosseini, A. *New discretization of total variation functional for image processing tasks*, Signal Process. Image Commun. 78 (2019), 62–76.

14. Huang, J., Zhang, S. and Metaxas, D. *Efficient MR image reconstruction for compressed MR imaging,* Med. Image Anal. 15 (2011), 670–679.

15. Knoll, F., Bredies, K., Pock, T. and Stollberger, R. *Second order total generalized variation (TGV) for MRI,* Magn. Reson. Med. 65 (2011), 480–491.

16. Kou, G., Zhao, Y., Peng, Y. and Shi, Y. *A fast alternating direction method for TVL1-L2 signal reconstruction from partial Fourier data,* IEEE J-STSP. 4 (2010), 288–297.

17. Lore, K.G., Akintayo, A. and Sarkar, S. *LLNEt: A deep autoencoder approach to natural low-light image enhancement*, Pattern Recognit, 61 (2015), 650–662.

18. Lustig, M. *Michael Lustig homepage,* https://people.eecs.berkeley.edu/~mlustig/index.html, Online, accessed February 2019.

19. Lustig, M., Donoho, D. and Pauly, J. *Sparse MRI: The application of compressed sensing for rapid MR imaging,* Magn. Reson. Med. 58 (2007), 1182–1195.

20. Ma, S., Yin, W., Zhang, Y. and Chakraborty, A. *An efficient algorithm for compressed MR imaging using total variation and wavelets,* IEEE Conference on Computer Vision and Pattern Recognition, CVPR (2008), 1–8.

21. Perona, P. and Malik, J. *Scale-space and edge detection using anisotropic diffusion*, IEEE Trans. Pattern Anal. Mach. Intell. 12 (1990), 629–639.

22. Ravishankar, S. and Bresler, Y. *MR image reconstruction from highly undersampled k-space data by dictionary learning,* IEEE Trans. Med. Imaging, 30 (2011), 1029–1041.

23. Ravishankar, S. and Bresler, Y. *Efficient blind compressed sensing using sparsifying transforms with convergence guarantees and application to magnetic resonance imaging,* SIAM J. Imaging Sci. 8 (2015), 2519–2557.

24. Rudin, L., Osher, S. and Fatemi, E. *Nonlinear total variation based noise removal algorithms*, Physica D. 60 (1992), 259–268.

25. Wang, N., Tao, D., Gao, X., Li, X. and Li, J. *A comprehensive survey to face hallucination*, Int. J. Comput. Vis. 106 (2014), 9–30.

26. Weickert, J. Anisotropic diffusion in image processing, Teubner, Stuttgart, 1998.