Mathematical Biosciences and Engineering

Research article

A new difference of anisotropic and isotropic total variation regularization method for image restoration

Benxin Zhang[1]**, Xiaolong Wang**[1]**, Yi Li**[1] **and Zhibin Zhu**[2,*]

[1] School of Electronic Engineering and Automation, Guilin University of Electronic Technology, Guilin 541004, China

[2] School of Mathematics and Computing Science, Guilin University of Electronic Technology, Guilin 541004, China

* **Correspondence:** Email: optimization_zhu@163.com.

Abstract: Total variation (TV) regularizer has diffusely emerged in image processing. In this paper, we propose a new nonconvex total variation regularization method based on the generalized Fischer-Burmeister function for image restoration. Since our model is nonconvex and nonsmooth, the specific difference of convex algorithms (DCA) are presented, in which the subproblem can be minimized by the alternating direction method of multipliers (ADMM). The algorithms have a low computational complexity in each iteration. Experiment results including image denoising and magnetic resonance imaging demonstrate that the proposed models produce more preferable results compared with state-of-the-art methods.

Keywords: image restoration; TV; Fischer-Burmeister function; DCA; ADMM

1. Introduction

Total variation [1] regularization method is widely used in image processing, such as image reconstruction and denoising. Generally, there are two types: isotropic and anisotropic. TV has a strong edge preserving ability and can process piecewise constant images. However, the image obtained by the TV often suffers from staircasing artifacts. Therefore, in order to suppress artifacts and improve image quality, TV variants have been studied by many scholars [2–6].

TV consists of the L_1 norm and gradient operator. From the factor of statistics, the L_1 norm may have deviation estimations for larger coefficients, where it will lose the oracle property [7]. Instead of the L_1 norm, the L_0 norm is the appropriate choice to ensure sparseness. However, the minimization of L_0 is a combination optimization problem and NP hard. Then, several penalty functions that integrate the benefits of L_1 and L_0 norms while circumventing their limitations have been suggested, such as

smoothly clipped absolute deviation [8], capped-L_1 [9], transformed-L_1 [10], and $L_1 - L_2$ [11]. Numerous algorithms employing convex norms on gradients have also proposed, including the alternating minimization algorithm [12] and the augmented Lagrange multiplier algorithm [13]. The alternating direction method of multipliers (ADMM) [14] and primal-dual algorithm [15] are two benchmark methods. In fact, the primal-dual hybrid gradient method introduced by Chambolle and Pock in [15] is equivalent to the preconditioned ADMM. Additionally, the widely used split Bregman method [16] can be considered as an instance of ADMM with two variables. Additionally, there are also many algorithms for solving nonconvex models. As variational models are inherently nonconvex, obtaining the solution for global optimization is difficult. Some approaches, such as the proximal alternating linearized minimization method [17] and the difference-of-convex algorithm (DCA) [18], typically converge to local minimizers. In [19], DCA was firstly created by Dinh et al. for solving the DC programming. Because of its simplicity and efficiency, it has been widely used and extensively studied.

DC programming and DCA are the natural extension from modern convex analysis to nonconvex analysis, broad enough to include most real world nonconvex optimizations. During the last decades, these theoretical and algorithmic tools have been greatly enriched. He et al. [20] proposed a unified Douglas-Rachford algorithm for generalized DC programming. Niu et al. [21] developed a refined inertial DC algorithm for DC programming. Artacho et al. [22, 23] investigated the line search idea in a DC program with linear constraint. Wen et al. [24] used the Nesterov acceleration technique for a special class of DC program. Moreover, DC programming and DCA have always been popular for nonconvex models in image processing, such as image restoration, image reconstruction, and compressive sensing. In [25, 26], Lou et al. proved that $L_1 - L_2$ is closer to L_0 than L_1 and developed a weighted difference of anisotropic and isotropic TV as a regularization for image denoising, deblurring, segmentation and magnetic resonance imaging (MRI). Its solution was obtained by DCA and the split Bregman. In [27], Li et al. introduced a multiplicative noise removal model, whose objective function can be written as the difference of two convex functions. DCA with a primal-dual hybrid gradient method was used for solving nonconvex programming. Sun et al. [28] studied the DCA to address the Log-norm TV image restoration problem and ADMM was used to solve the resulting subproblem. In [29], for impulsive noise removal, DCA with adaptive proximal parameters was proposed. DCA is a simple descent algorithm. Interested readers are encouraged to consult these state of the art developments in [30].

The contributions of this study are as follows. First, motivated by the idea that $L_1 - L_2$ can be considered as a Fischer-Burmeister (FB) function, we propose a new nonconvex TV regularization term based on a general FB function. Second, two efficient DCAs are designed for solving nonconvex optimization problems. The proposed DCA needs to solve the TV type of subproblem, which can be efficiently handled by ADMM. Last, experiments on image denoising and MRI reconstruction show that the proposed methods improve on the classical TV model and are compared with state-of-the-art methods. Especially, for MRI, the new model can successfully reconstruct the Shepp-Logan image using merely 7 radial lines.

The rest of this paper is organized as follows. We show our new models in Section 2. Section 3 presents the DCA with ADMM for solving the new nonconvex models. Section 4 shows the effectiveness of the new models and algorithms through the presentation of experimental results. The conclusion is provided in the last section.

2. New models

The problem of restoring image with additive noise is the following:

$$f = Au + e, \tag{2.1}$$

where $f \in R^m$ is noise image, $A \in R^{m \times n}$ is a linear operator, $u \in R^n$ is original image and $e \in R^m$ is Gaussian noise. To simplify the presentation, we express both the original and observation images or data as vectors.

Our aim is to restore the initial image u from the noisy data f. Image restoration is a typical inverse problem, and it's impossible to obtain the original image u from f by just the direct inversion of (2.1). A classic method is variational regularization for addressing the ill-posed problem. This approach can make the solutions stable. Then, for the Gaussian noise removal, the famous TV model is as follows:

$$\min_u \|u\|_{TV}, \ s.t. \ Au = f,$$

where $\|\cdot\|$ represents the Euclidean norm. As discussed above, TV can be classified into two categories: isotropic and anisotropic:

$$\|Du\|_{2,1} = \| \sqrt{|D_x u|^2 + |D_y u|^2} \|_1, \quad \|Du\|_1 = \|D_x u\|_1 + \|D_y u\|_1,$$

where D is the gradient operator, D_x and D_y are the horizontal and vertical derivative operators, respectively, and $\|\cdot\|_1$ means the L_1 norm.

Illuminated by $L_1 - L_2$ minimization in compression sensing [11], Lou et al. [26] proposed the well-known $L_1 - \gamma L_2$ regularization function, which is the coupling of anisotropic and isotropic TV:

$$J(u) = \|Du\|_1 - \gamma \|Du\|_{2,1} = \|D_x u\|_1 + \|D_y u\|_1 - \gamma \| \sqrt{|D_x u|^2 + |D_y u|^2} \|_1,$$

where γ is an element of the interval $(0,1]$. In [31], the authors find that $J(u)$ can be considered as the application of the FB function to the gradient of TV (FBTV), where the scalar-valued FB function is defined as

$$\phi(v_1, v_2) = v_1 + v_2 - \gamma \sqrt{v_1^2 + v_2^2}, \quad \gamma \in (0, 1].$$

FBTV considers the non-sparse gradient vectors along the boundaries in the image. The effectiveness of FBTV was demonstrated to surpass that of either anisotropic or isotropic TV.

Motivated by $J(u)$ can be described as a FB function and $L_1 - L_2$ is closer to L_0 than L_1, we propose a new difference regularization method by using the following generalized FB (GFB) function:

$$\psi(v_1, v_2) = v_1 + v_2 - \sqrt{(1 - \theta)(v_1^2 + v_2^2) + \theta(v_1 - v_2)^2}, \quad \theta \in [0, 1).$$

Now, we present the nonconvex variational regularization models, named GFBTV-C, as

$$\min_u \|D_x u\|_1 + \|D_y u\|_1 - \| \sqrt{(1 - \theta)(|D_x u|^2 + |D_y u|^2) + \theta(|D_x u| - |D_y u|)^2} \|_1, \tag{2.2}$$
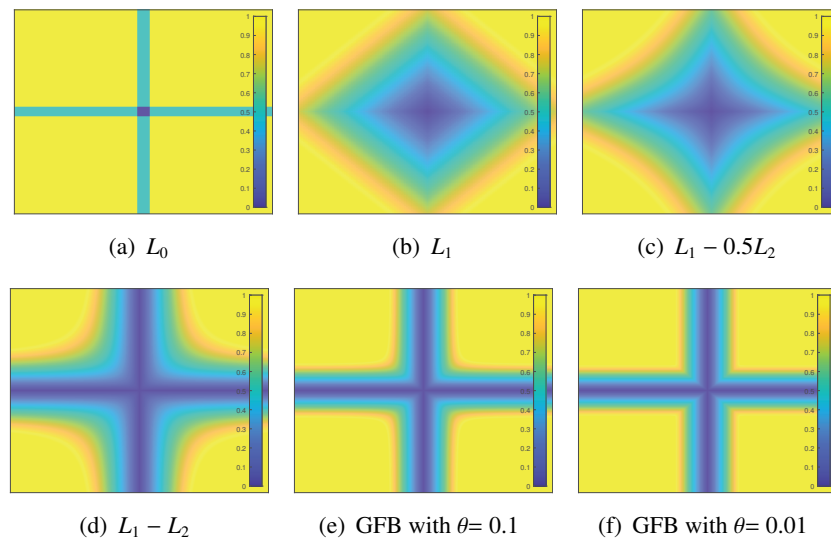$$s.t. \ Au = f.$$

Figure 1. Level curves of different metrics. GFB function with θ=0.1 and 0.01 is closer to L_0 than L_1, $L_1 - 0.5L_2$ and $L_1 - L_2$.

and GFBTV as

$$\min_u \ \|D_xu\|_1 + \|D_yu\|\|_1 - \|\sqrt{(1-\theta)(|D_xu|^2 + |D_yu|^2) + \theta(|D_xu| - |D_yu|)^2}\|_1 + \frac{\alpha}{2}\|Au - f\|^2, \qquad (2.3)$$

where $\alpha > 0$ is the regularization parameter. It is observed that $\psi(v_1, v_2)$ takes into account other gradient directions, not just the horizontal or vertical. This allows it to regulate the impact of anisotropic TV for detecting sharp edges while reducing the blocky artifacts. To better understand the different metrics, we plot the level curves corresponding to L_0, L_1, $L_1 - 0.5L_2$, $L_1 - L_2$ and GFB with θ=0.1 and 0.01 in Figure 1. From the figure, it can observe that the contour lines of GFB are bending more inward and closer to L_0. This illustrates that GFB can enhance sparsity, and θ acts like a parameter controlling to what extent.

3. Numerical algorithms

Though $\|\sqrt{(1-\theta)(|D_xu|^2 + |D_yu|^2) + \theta(|D_xu| - |D_yu|)^2}\|_1$ is not convex, DCA can also be used to easily solve the GFBTV-C (2.2) and GFBTV (2.3) models. In the following, we apply the DCA with ADMM to handle the two nonsmooth and nonconvex models.

3.1. DCA for solving the model GFBTV-C

First, in order to use the DCA, we separate into (2.2) as follows:

$$\min_u \Gamma(u) = \Phi(u) - \Psi(u), \ s.t. \ Au = f, \qquad (3.1)$$

where

$$\begin{cases} \Phi(u) = \|D_xu\|_1 + \|D_yu\|\|_1, \\ \Psi(u) = \|\sqrt{(1-\theta)(|D_xu|^2 + |D_yu|^2) + \theta(|D_xu| - |D_yu|)^2}\|_1. \end{cases}$$

DCA optimizes (3.1) via linearizing $\Psi(u)$:

$$u_{k+1} = \arg\min_u \{\Phi(u) - (\Psi(u_k) + \langle q^k, D(u - u_k)\rangle),\ s.t.\ Au = f\}, \tag{3.2}$$

where

$$q^k = (q_x^k, q_y^k) = \frac{(D_x u_k - \theta D_y u_k, D_y u_k - \theta D_x u_k)}{\sqrt{(1 - \theta)(|D_x u_k|^2 + |D_y u_k|^2) + \theta(|D_x u_k| - |D_y u_k|)^2}}.$$

If the denominator is zero, the corresponding value is set as zero. Therefore, by the form of ADMM, the convex subproblem (3.2) could be rewritten as the unconstrained problems:

$$u_{n+1} = \arg\min_u \{\Phi(u) - (\Psi(u_k) + \langle q^k, D(u - u_k)\rangle) + \frac{\mu}{2}\|Au - z_n\|^2\}, \tag{3.3}$$

$$z_{n+1} = z_n + f - Au_{n+1}.$$

where z is a Lagrange multiplier which enforces $Au = f$ and the parameter μ is the penalty parameter.

The subproblem (3.3) equals to solve the TV problem. Therefore, ADMM is also used to find the solution. In order to utilize ADMM, two auxiliary variables, d_x and d_y, are introduced to rephrase the original minimization problem (3.3) as a constrained problem:

$$\min_{u,d_x,d_y} \|d_x\|_1 + \|d_y\|\|_1 - (q_x^k d_x + q_y^k d_y) + \frac{\mu}{2}\|Au - z_n\|^2,$$
$$s.t.\ d_x = D_x u, d_y = D_y u. \tag{3.4}$$

As per the classical optimization theory, the augmented Lagrangian function for Eq (3.4) can be expressed as follows:

$$\begin{aligned}\mathcal{L}(u, d_x, d_y, b_x, b_y) &= \|d_x\|_1 + \|d_y\|\|_1 - (q_x^k d_x + q_y^k d_y) + \frac{\mu}{2}\|Au - z_n\|^2 \\ &+ \frac{\lambda}{2}\|d_x - D_x u - b_x\|^2 + \frac{\lambda}{2}\|d_y - D_y u - b_y\|^2,\end{aligned} \tag{3.5}$$

where b_x, b_y are the Lagrange multipliers and $\lambda > 0$ is the penalty parameter.

Subsequently, our attention turns to obtain the solution for subproblems in (3.5). The u-subproblem can be written as follows:

$$u^{j+1} = \arg\min_u \mathcal{L}(u, d_x^j, d_y^j, b_x^j, b_y^j). \tag{3.6}$$

Using the first-order optimality condition of Eq (3.6), we have

$$(\mu A^T A - \lambda D^T D)u^{j+1} = \lambda D_x^T(d_x^j - b_x^j) + \lambda D_y^T(d_y^j - b_y^j) + \mu A^T z_n. \tag{3.7}$$

In the problem of reconstructing undersampled magnetic resonance imaging (MRI) using compressed sensing [32], the composite $A = RF$ is formed by combining a sampling operator R and a Fourier transform operator F. When periodic boundary conditions are applied, both matrices $D^T D$ and $A^T A$ become block circulant with circulant blocks. As a result, the coefficient matrix $\mu A^T A - \lambda D^T D$ can be diagonalized using a discrete Fourier transform. This means that Eq (3.7) can be efficiently solved by a fast Fourier transform (FFT):

$$u^{j+1} = \mathcal{F}^{-1}\left(\frac{\mathcal{F}(\lambda D_x^T(d_x^j - b_x^j)) + \mathcal{F}(\lambda D_y^T(d_x^j - b_x^j)) + \mu\mathcal{F}(A)^* \odot \mathcal{F}(z_n)}{\mu\mathcal{F}(A)^* \odot \mathcal{F}(A) - \lambda\mathcal{F}(D)^* \odot \mathcal{F}(D)}\right).$$

Here, $\mathcal{F}(\cdot)$ and $\mathcal{F}^{-1}(\cdot)$ refer to the FFT and inverse FFT operations, respectively. The symbol " $*$ " represents a complex conjugation, while " $\odot$ " denotes an elementwise multiplication.

For d_x, d_y-subproblem in (3.5), the solution can be updated via soft shrinkage

$$d_x^{j+1} = \text{shrink}(D_x u^{j+1} + b_x^j + q_x^k/\lambda, 1/\lambda), \tag{3.8}$$

$$d_y^{j+1} = \text{shrink}(D_y u^{j+1} + b_y^j + q_y^k/\lambda, 1/\lambda), \tag{3.9}$$

where

$$\text{shrink}(s, \mu) = \text{sign} \max\{|s| - \mu, 0\},$$

and sign as the signum function. Additionally, the Lagrange multipliers are updated as follows:

$$b_x^{j+1} = b_x^j + D_x u^{j+1} - d_x^{j+1}, \tag{3.10}$$

$$b_y^{j+1} = b_y^j + D_y u^{j+1} - d_y^{j+1}. \tag{3.11}$$

At last, we present the DCA with ADMM for solving the proposed model GFBTV-C in Algorithm 1.

Algorithm 1 DCA-ADMM for solving the model GFBTV-C (2.2).

Set $u_0 = q_x^0 = q_y^0 = 0$, $z_0 = f$, $\lambda_0 = 0$, MaxDCA, k_{max}, j_{max}
For $k = 0, 1, 2, \cdots$, MaxDCA, $b_x^0 = b_y^0 = 0$,
  For $n = 0, 1, 2, \cdots, n_{max}$
    For $j = 0, 1, 2, \cdots, j_{max}$
      Calculate u^{j+1} by (3.7),
      Calculate d_x^{j+1} by (3.8),
      Calculate d_y^{j+1} by (3.9),
      Calculate b_x^{j+1} by (3.10),
      Calculate b_y^{j+1} by (3.11).
    End For

$$u_{n+1} = u^{j_{max}},$$

$$z_{n+1} = z_n + f - A u_{n+1}.$$

  End For

$$u_{k+1} = u^{n_{max}},$$

$$(q_x^{k+1}, q_y^{k+1}) = \frac{(D_x u_{k+1} - \theta D_y u_{k+1}, D_y u_{k+1} - \theta D_x u_{k+1})}{\sqrt{(1 - \theta)(|D_x u_{k+1}|^2 + |D_y u_{k+1}|^2) + \theta(|D_x u_{k+1}| - |D_y u_{k+1}|)^2}}.$$

End For

3.2. DCA for solving the model GFBTV

For the corresponding unconstrained problem (2.3), the DCA subproblem is expressed as

$$u_{k+1} = \arg \min_u \{\Phi(u) - (\Psi(u_k) + \langle q^k, D(u - u_k)\rangle) + \frac{\alpha}{2}\|Au - f\|^2\}.$$

Similar to (3.5), we have

$$\min_{u,d_x,d_y} \|d_x\|_1 + \|d_y\|\|_1 - (q_x^k d_x + q_y^k d_y) + \frac{\alpha}{2}\|Au - f\|^2,$$
$$s.t.\ d_x = D_x u, d_y = D_y u. \tag{3.12}$$

The augmented Lagrangian function for Eq (3.12) can be written as

$$\mathcal{L}(u, d_x, d_y, b_x, b_y) = \|d_x\|_1 + \|d_y\|\|_1 - (q_x^k d_x + q_y^k d_y) + \frac{\alpha}{2}\|Au - f\|^2$$
$$+ \frac{\lambda}{2}\|d_x - D_x u - b_x\|^2 + \frac{\lambda}{2}\|d_y - D_y u - b_y\|^2. \tag{3.13}$$

The next step is to express the u-subproblem as

$$u^{j+1} = \arg\min_u \mathcal{L}(u, d_x^j, d_y^j, b_x^j, b_y^j). \tag{3.14}$$

Then, using the first-order optimality condition of Eq (3.14), we get

$$(\alpha A^T A - \lambda D^T D)u^{j+1} = \lambda D_x^T(d_x^j - b_x^j) + \lambda D_y^T(d_y^j - b_y^j) + \alpha A^T f. \tag{3.15}$$

Moreover, if $D^T D$ and $A^T A$ have special structure, Eq (3.15) can be solved via FFT as follows:

$$u^{j+1} = \mathcal{F}^{-1}\left(\frac{\mathcal{F}(\lambda D_x^T(d_x^j - b_x^j)) + \mathcal{F}(\lambda D_y^T(d_x^j - b_x^j)) + \alpha \mathcal{F}(A)^* \odot \mathcal{F}(f)}{\alpha \mathcal{F}(A)^* \odot \mathcal{F}(A) - \lambda \mathcal{F}(D)^* \odot \mathcal{F}(D)}\right). \tag{3.16}$$

For d_x, d_y-subproblem in (3.13), the solution can be updated via soft shrinkage

$$d_x^{j+1} = \text{shrink}(D_x u^{j+1} + b_x^j + q_x^k/\lambda, 1/\lambda), \tag{3.17}$$

$$d_y^{j+1} = \text{shrink}(D_y u^{j+1} + b_y^j + q_y^k/\lambda, 1/\lambda), \tag{3.18}$$

and the Lagrange multipliers are updated as follows:

$$b_x^{j+1} = b_x^j + D_x u^{j+1} - d_x^{j+1}, \tag{3.19}$$

$$b_y^{j+1} = b_y^j + D_y u^{j+1} - d_y^{j+1}. \tag{3.20}$$

Algorithm 2 is presented to demonstrate the DCA for resolving the proposed model GFBTV. To tackle the unconstrained problem (2.3), Algorithm 2 is nearly identical to Algorithm 1, but differs by including an extra update on z.

4. Experiments

We present the experimental results of model GFBTV-C (2.2) utilizing Algorithm 1 for MRI and model GFBTV (2.3) employing Algorithm 2 for image denoising. From Figure 1, we find that θ=0.01 and 0.1 can approximate the L_0 norm perfectly. Hence, we choose the control parameter θ=0.01 for image denoising and θ=0.1 for MRI naturally. All experiments were executed on a desktop computer with a 2.9 GHz processor and 16GB RAM. The peak signal-to-noise ratio (PSNR) is used to evaluate the restored quality, which is defined as

$$PSNR = 20 * log_{10}\frac{255}{MSE}dB.$$

where MSE represents the mean square error between u_k and u_{ori}, u_k is the recovered image, and u_{ori} represents the original image. Furthermore, it is worth noting that the structure similarity (SSIM) [33] has also been reported. It knows that as the PSNR and SSIM values increase, the quality of the reconstructed image improves.

Algorithm 2 DCA-ADMM for solving the model GFBTV.

Set $u_0 = q_x^0 = q_y^0 = 0$, $z_0 = f$, $\lambda_0 = 0$, MaxDCA, j_{max},

For $k = 0, 1, 2, \cdots$, MaxDCA, $b_x^0 = b_y^0 = 0$,

    For $j = 0, 1, 2, \cdots, j_{max}$

        Calculate u^{j+1} by (3.15),

        Calculate d_x^{j+1} by (3.17),

        Calculate d_y^{j+1} by (3.18),

        Calculate b_x^{j+1} by (3.19),

        Calculate b_y^{j+1} by (3.20).

    End For

$$u_{k+1} = u^{j_{max}},$$

$$(q_x^{k+1}, q_y^{k+1}) = \frac{(D_x u_{k+1} - \theta D_y u_{k+1}, D_y u_{k+1} - \theta D_x u_{k+1})}{\sqrt{(1 - \theta)(|D_x u_{k+1}|^2 + |D_y u_{k+1}|^2) + \theta(|D_x u_{k+1}| - |D_y u_{k+1}|)^2}}.$$

  End For

Table 1. Average results of six models on PSNR and SSIM values for image denoising.

Image	σ	DnCNN		TV		$L_1 - 0.5L_2$		$L_1 - L_2$		TTV		GFBTV	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Pascal	0.05	34.39	0.9048	31.98	0.8772	33.18	0.8953	32.88	0.8829	33.01	0.8930	33.48	0.8976
	0.1	30.79	0.8319	27.62	0.7995	29.04	0.8231	29.53	0.8283	29.84	0.8254	29.68	0.8290

4.1. Image denoising

In this subsection, we will discuss the issue of image denoising, in which the operator A is set to an identity matrix for the nonconvex model GFBTV. We compare it with several well-known variational models and a deep learning-based approach named DnCNN [34]. The first variational model we consider is the TV model, which is also known as the classic ROF model [1]. The second model is $L_1 - \gamma L_2$ with γ equaling 0.5 or 1, as discussed in [26]. The last model is the transform total variation (TTV) in [35], and DCA is formulated for the TTV minimization problem. We can efficiently solve the new model GFBTV using Algorithm 2, where u_{k+1} can be calculated by (3.16) with $A = I$.

For this experiment, 100 images were randomly selected from the Pascal dataset as the test images. Both of these images were subject to degradation caused by additive Gaussian noise with a mean of zero and a standard deviation of σ. The values of σ chosen for the experiment are 0.05 and 0.1. The penalty parameter μ was set to 5 for σ=0.1 and 15 for σ=0.05 during testing. For the proposed model, we set $\theta = 0.01$. The default parameter settings are the same as the references [26, 35] for a fair comparison. We initialize $u_0 = f$, and the maximum number of DCA is 2.

The average comparison results are reported in Table 1. The PSNR and SSIM values of the different algorithms for various levels of σ are provided. It is shown from Table 1 that DnCNN is superior to the gradient models in terms of PSNR and SSIM values. Although our new approach has slightly lower PSNR and SSIM values compared to DnCNN, it outperforms TV based models for various noise levels about SSIM. The denoising results of the six models are presented in Figures 2–5. All of the methods effectively remove noise as shown in the four figures and some fine structures and details are easy to see from the zoomed regions. DnCNN demonstrates the highest performance regarding PSNR among the
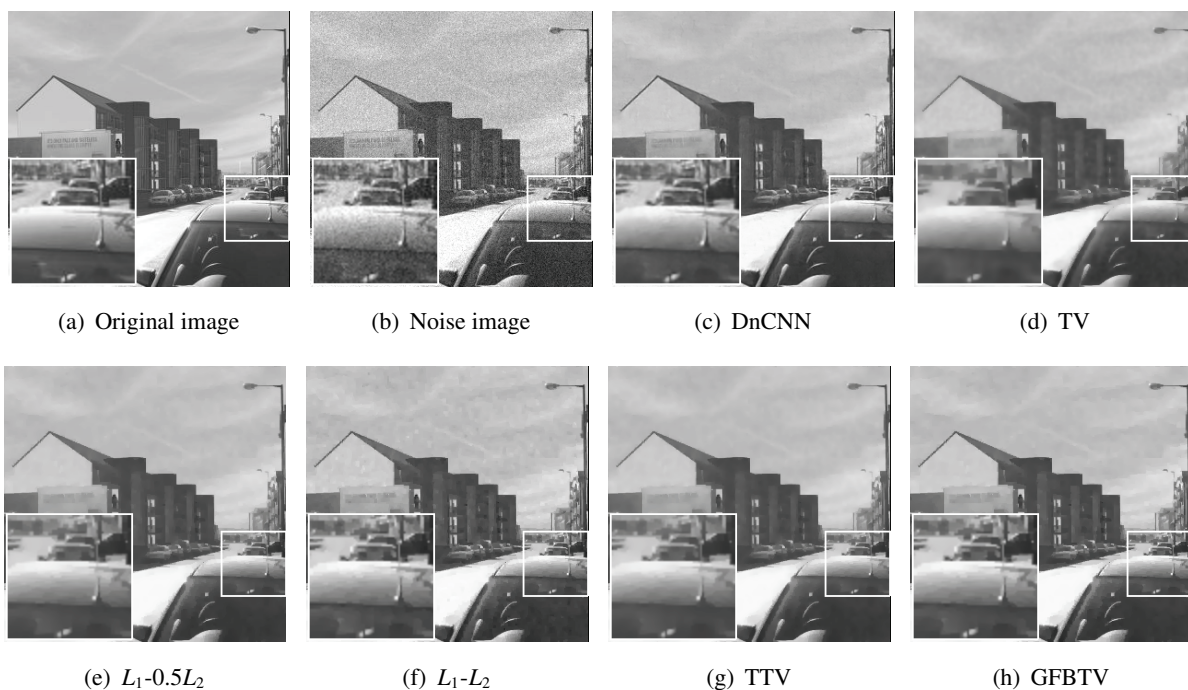
Figure 2. (a) Original image, (b) noisy image with σ=0.05, (c) DnCNN, PSNR=32.70, SSIM=0.9158, (d) TV, PSNR=29.06, SSIM=0.9080, (e) L_1-0.5L_2, PSNR=30.75, SSIM=0.9080, (f) L_1-L_2, PSNR=31.46, SSIM=0.9034, (g) TTV, PSNR=30.72, SSIM=0.9070, (h) GFBTV, PSNR=31.57, SSIM=0.9153.
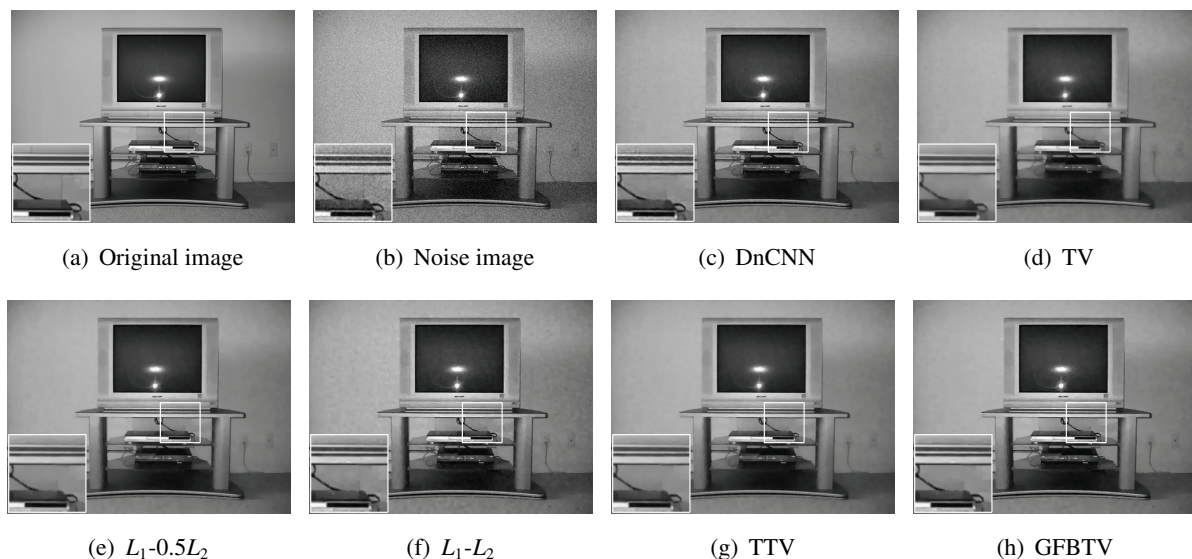

Figure 3. (a) Original image, (b) noisy image with σ=0.05, (c) DnCNN, PSNR=35.80, SSIM=0.9341, (d) TV, PSNR = 31.38, SSIM=0.9193, (e) L_1-0.5L_2, PSNR = 33.77, SSIM=0.9394, (f) L_1-L_2, PSNR=34.62, SSIM=0.9280, (g) TTV, PSNR=33.31, SSIM=0.9358, (h) GFBTV, PSNR=34.94, SSIM=0.9425.

(a) Original image (b) Noise image (c) DnCNN (d) TV
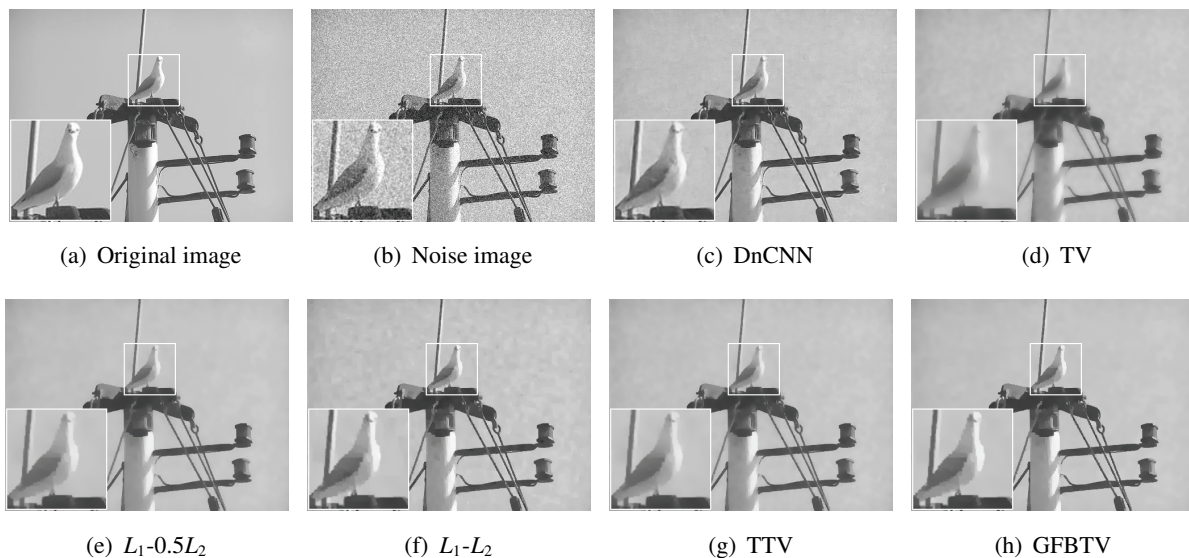
(e) L_1-0.5L_2 (f) L_1-L_2 (g) TTV (h) GFBTV

Figure 4. (a) Original image, (b) noisy image with σ=0.1, (c) DnCNN, PSNR=31.23, SSIM=0.8926, (d) TV, PSNR=27.28, SSIM=0.9079, (e) L_1-0.5L_2, PSNR=29.19, SSIM=0.9372, (f) L_1-L_2, PSNR=29.63, SSIM=0.9061, (g) TTV, PSNR=28.94, SSIM=0.9349, (h) GFBTV, PSNR=30.14, SSIM=0.9439.

(a) Original image (b) Noise image (c) DnCNN (d) TV

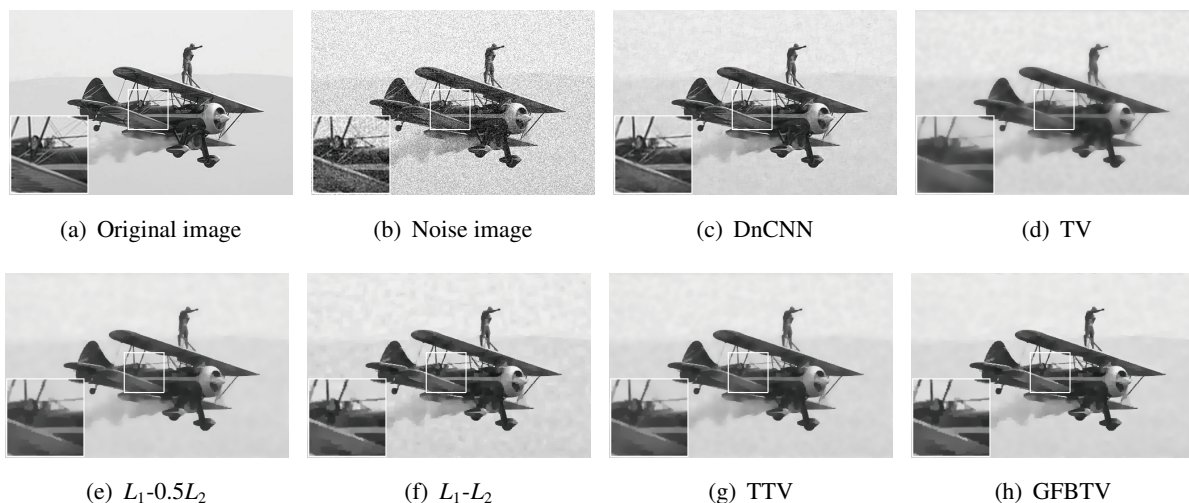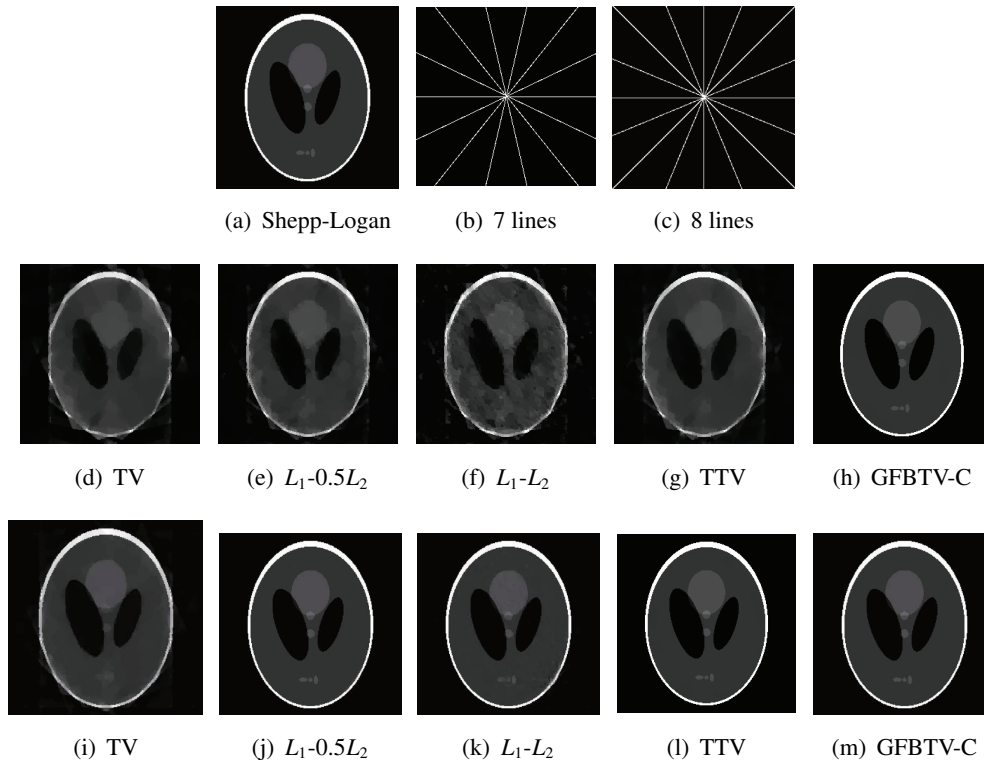(e) L_1-0.5L_2 (f) L_1-L_2 (g) TTV (h) GFBTV

Figure 5. (a) Original image, (b) noisy image with σ=0.1, (c) DnCNN, PSNR=31.60, SSIM=0.8902, (d) TV, PSNR=27.28, SSIM=0.8983, (e) L_1-0.5L_2, PSNR=28.84, SSIM=0.9207, (f) L_1-L_2, PSNR=29.40, SSIM=0.9016, (g) TTV, PSNR=28.61, SSIM=0.9171, (h) GFBTV, PSNR=29.67, SSIM=0.9266.

Table 2. Performance comparisons of the five models for MRI.

Image	Mask	TV			$L_1 - 0.5L_2$			$L_1 - L_2$			TTV			GFBTV-C		
		PSNR	SSIM	Time	PSNR	SSIM	Time	PSNR	SSIM	Time	PSNR	SSIM	Time	PSNR	SSIM	Time
Shepp-Logan	7 lines	18.44	0.4541	29.21	21.45	0.6624	34.45	20.66	0.5615	34.56	20.25	0.5643	33.81	42.17	0.9937	34.12
	8 lines	24.22	0.7148	28.87	48.82	0.9971	33.90	37.01	0.9702	34.46	53.83	0.9965	33.03	63.19	0.9995	34.87
Brain-1	Cartesian	32.61	0.9564	29.89	35.73	0.9454	34.20	35.37	0.9465	33.34	31.59	0.9425	34.51	36.65	0.9741	33.95
Brain-2	Random	30.49	0.8614	29.56	31.57	0.8661	34.96	30.99	0.8122	34.17	30.09	0.8728	33.18	32.42	0.8922	33.87

(a) Shepp-Logan (b) 7 lines (c) 8 lines

(d) TV (e) L_1-0.5L_2 (f) L_1-L_2 (g) TTV (h) GFBTV-C

(i) TV (j) L_1-0.5L_2 (k) L_1-L_2 (l) TTV (m) GFBTV-C

Figure 6. Reconstructed the Shepp-Logan phantom image using 7 lines and 8 lines.

compared approaches. However, when evaluating the four test images, our proposed method exhibits higher SSIM values than both DnCNN and other TV based models, and can reserve more details of the original images. It is also obvious that artifacts appear in the DnCNN, TV and $L_1 - L_2$ methods. Hence, the experimental results show that the new model can effectively reduce the blocky artifacts and obtain better visual effects.

4.2. Magnetic resonance imaging

In this subsection, we consider the problem of MR reconstruction. Compressive sensing based magnetic resonance imaging technology is widely used in images solving inverse problems. The sparse image reconstruction model recovers images from under-sampled k-space data, which effectively solves the problem of data acquisition efficiency in MRI domain. Consequently, we adopt the constrained formulation GFBTV-C (2.2), and Algorithm 1 is used for solving it. Additionally, we compare with TV, $L_1 - \gamma L_2$ with γ=0.5 or 1 in [26], and TTV [35]. In this test, θ is set as 0.1. The related parameters in the proposed method and competing methods are left to their default settings in [26, 35].

First, we consider the issue of MRI reconstruction utilizing a Shepp-Logan phantom from seven and eight radial projections. The results of various models in reconstructing the Shepp-Logan phantom are shown in Figure 6. Among these methods, only GFBTV-C can attain precise recovery using 7 lines, and the result is excellent. When it comes to 8 lines, GFBTV-C, TTV and $L_1 - 0.5L_2$ can achieve perfect reconstruction. TV and $L_1 - L_2$ methods have some artifacts in the reconstructed images.
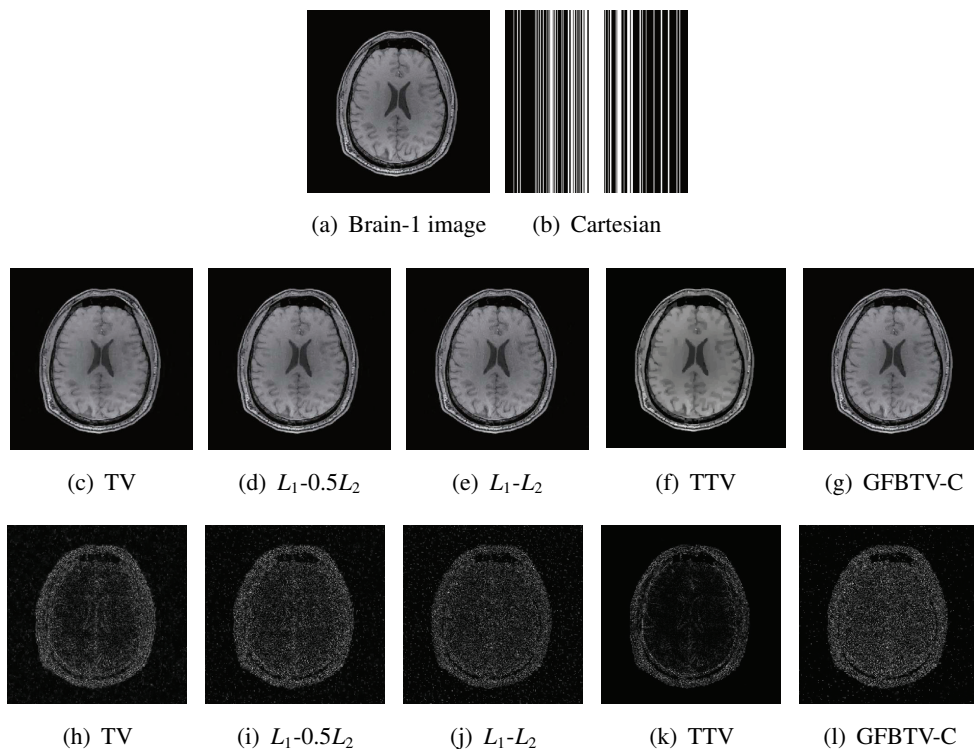

(a) Brain-1 image   (b) Cartesian

(c) TV   (d) L_1-$0.5L_2$ (e) L_1-L_2 (f) TTV   (g) GFBTV-C

(h) TV   (i) L_1-$0.5L_2$ (j) L_1-L_2 (k) TTV   (l) GFBTV-C

Figure 7. Original Brain-1 image, the reconstruction and residue images by different models.

Second, we use two brain images test as the test images: Brain-1 and Brain-2 with dimensions of 256×256. These are generated from the literatures [36, 37]. Then, we perform undersampling reconstruction under different sampling patterns. The k-space data for Brain-1 were randomly sampled using a Cartesian mask with 87 lines. For Brain-2, a variable density mask with a sampling rate of 30% was used to undersample the k-space data. The brain images reconstructed the residue images between the reconstruction images and original images by the five models, and are shown in Figures 7 and 8. All methods have better recovery effects. Visually, there is no significant difference. The residue images illustrate that the proposed method can visually gain better reconstructions. It appears that the TV model has the most residual images. From the residual images, it can be observed that TTV, $L_1 - 0.5L_2$, $L_1 - L_2$ methods have similar effects on the image details. Table 2 presents the numerical performance of the five compared models. From the table, we see that the TV model needs less CPU time (in second). TTV, $L_1 - 0.5L_2$, $L_1 - L_2$ and our GFBTV-C have similar time costs. Meanwhile, we know that GFBTV-C provides an improved reconstruction quality than TV, $L_1 - 0.5L_2$, and $L_1 - L_2$ as it has the highest PSNR and SSIM values. The PSNR values of the new method are improved by at least 0.82 dB, and the SSIM values around 0.02. As the sampling rate becomes smaller, the performance of

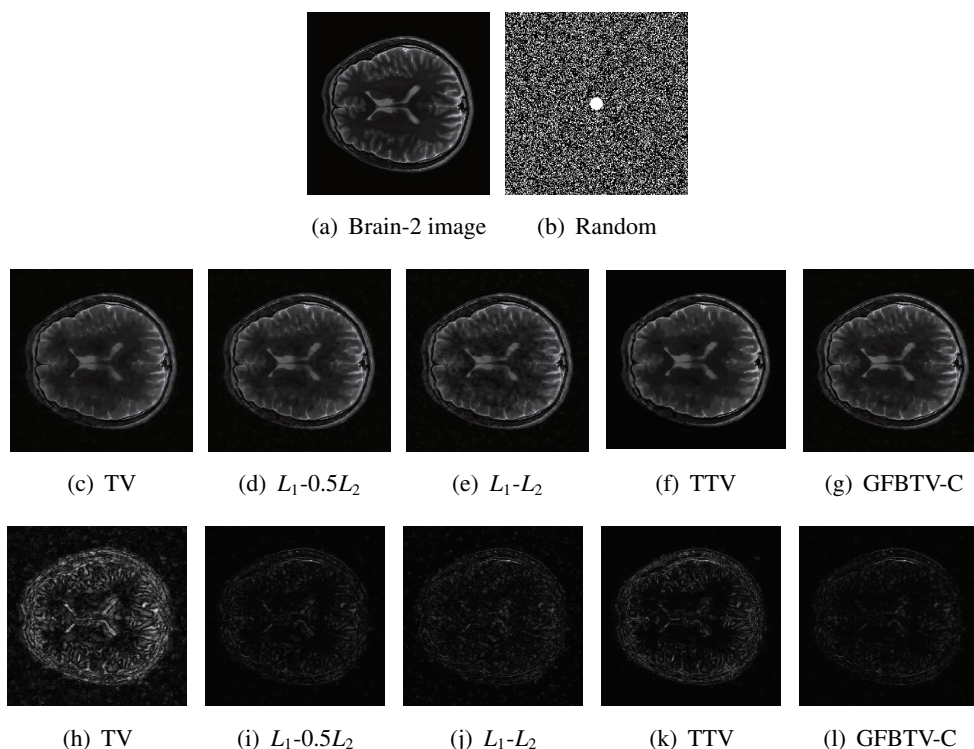the proposed model improves more significantly.

(a) Brain-2 image   (b) Random

(c) TV   (d) L_1-$0.5L_2$ (e) L_1-L_2 (f) TTV   (g) GFBTV-C

(h) TV   (i) L_1-$0.5L_2$ (j) L_1-L_2 (k) TTV   (l) GFBTV-C

Figure 8. Original Brain-2 image, the reconstruction and residue images by different models.

5. Conclusions

This paper introduces a novel nonconvex TV term that utilizes the GFB function to restore images corrupted by Gaussian noise. Additionally, we use DCA to solve the optimization problem. The results of our experiments on MRI and image denoising exhibit that our proposed models are capable of achieving better solutions in comparison to both convex models and other nonconvex models, and produce more preferable results.

Besides, we observe that the new regularization function has an improved to encourage sparsity compared to L_1, $L_1 - \gamma L_2$ norm, and approximates L_0 norm more closely. This property leads to a better solution. Furthermore, the proposed model succeeds in reconstructing the Shepp-Logan phantom image from merely 7 radial lines. Hence, the superiority of our approach is more apparent when the sampling rate is relatively low or the available measurements are limited. The proposed model in this paper is mainly developed for image denoise and MRI, in the future, we will intend to extend the new method to computed tomography, Parallel MRI, electrical impedance tomography and so on. It is also essential to highlight that the convergence of our algorithm remains an open challenge and will be addressed in the future work.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgement

Conflict of interest

The authors declare no conflict of interest.

References

1. L. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms, *Phys. D*, **60** (1992), 259–268. https://doi.org/10.1016/0167-2789(92)90242-F

2. K. Bredies, K. Kunisch, T. Pock, Total generalized variation, *SIAM J. Imaging Sci.*, **3** (2010), 492–526. https://doi.org/10.1137/090769521

3. L. Condat, Discrete total variation: New definition and minimization, *SIAM J. Imaging Sci.*, **10** (2017), 1258–1290. https://doi.org/10.1137/16M1075247

4. Z. Jia, M. K. Ng, W. Wang, Color image restoration by saturation-value total variation, *SIAM J. Imaging Sci.*, **12** (2019), 972–1000. https://doi.org/10.1137/16M1075247

5. S. Pan, Q. Dai, H. Chen, Global optimality analysis and solution of the total variation signal denoising model, *Math. Biosci. Eng.*, **20** (2023), 6932–6946. https://doi.org/10.3934/mbe.2023299

6. D. Xiao, J. Li, R. Zhao, S. Qi, Y. Kang, Iterative CT reconstruction based on ADMM using shearlet sparse regularization, *Math. Biosci. Eng.*, **19** (2022), 11840–11853. https://doi.org/10.3934/mbe.2022552

7. R. Tibshirani, Regression shrinkage and selection via the lasso, *J. R. Stat. Soc. B*, **58** (1996), 267–288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

8. J. Fan, H. Peng, Nonconcave penalized likelihood with a diverging number of parameters, *Ann. Stat.*, **32** (2004), 928–961. https://doi.org/10.1214/009053604000000256

9. T. Zhang, Analysis of multi-stage convex relaxation for sparse regularization, *J. Mach. Learn. Res.*, **11** (2010), 1081–1107.

10. S. Zhang, J. Xin, Minimization of transformed L1 penalty: theory, difference of convex function algorithm, and robust application in compressed sensing, *Math. Program.*, **169** (2018), 307–336. https://doi.org/10.1007/s10107-018-1236-x

11. Y. Lou, P. Yin, Q. He, J. Xin, Computing sparse representation in a highly coherent dictionary based on difference of L1 and L2, *J. Sci. Comput.*, **64** (2015), 178–196. https://doi.org/10.1007/s10915-014-9930-1

12. Y. Wang, J. Yang, W. Yin, Y. Zhang, A new alternating minimization algorithm for total variation image reconstruction, *SIAM J. Imaging Sci.*, **1** (2008), 248–272. https://doi.org/10.1137/080724265

13. M. Fortin, R. Glowinski, On decomposition-coordination methods using an augmented lagrangian, *Stud. Math. Appl.*, **15** (1983), 97–146. https://doi.org/10.1016/S0168-2024(08)70028-6

14. S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, Distributed optimization and statistical learning via the alternating direction method of multipliers, *Found. Trends Mach. Learn.*, **3** (2011), 1–122. https://doi.org/10.1561/2200000016

15. A. Chambolle, T. Pock, A first-order primal-dual algorithm for convex problems with applications to imaging, *J. Math. Imaging Vision*, **40** (2011), 120–145. https://doi.org/10.1007/s10851-010-0251-1

16. T. Goldstein, S. Osher, The split Bregman method for L_1-regularized problems, *SIAM J. Imaging Sci.*, **2** (2009), 323–343. https://doi.org/10.1137/080725891

17. J. Bolte, S. Sabach, M. Teboulle, Proximal alternating linearized minimization for nonconvex and nonsmooth problems, *Math. Program.*, **146** (2014), 459–494. https://doi.org/10.1007/s10107-013-0701-9

18. H. L. Thi, T. P. Dinh, DC programming and DCA: thirty years of developments, *Math. Program.*, **169** (2018), 5–68. https://doi.org/10.1007/s10107-018-1235-y

19. P. D. Tao, Algorithms for solving a class of nonconvex optimization problems, methods of subgradients, in *North-Holland Mathematics Studies*, **129** (1986), 249–271. https://doi.org/10.1016/S0304-0208(08)72402-2

20. C. Chuang, H. He, Z. Zhang, A unified Douglas-Rachford algorithm for generalized DC programming, *J. Global Optim.*, **82** (2022), 331–349. https://doi.org/10.1007/s10898-021-01079-y

21. Y. You, Y. Niu, A refined inertial DC algorithm for DC programming, *Optim. Eng.*, **24** (2023), 65–91. https://doi.org/10.1007/s11081-022-09716-5

22. F. J. Aragon-Artacho, R. Campoy P. T. Vuong, The boosted DC algorithm for linearly constrained DC programming, *Set-Valued Var. Anal.*, **30** (2022), 1265–1289. https://doi.org/10.1007/s11228-022-00656-x

23. F. J. Aragon-Artacho, M. T. Fleming, P. T. Vuong, Accelerating the DC algorithm for smooth functions, *Math. Program.*, **169** (2018), 95–118. https://doi.org/10.1007/s10107-017-1180-1

24. B. Wen, X. Chen, T. K. Pong, A proximal diference-of-convex algorithm with extrapolation, *Comput. Optim. Appl.*, **69** (2018), 297–324. https://doi.org/10.1007/s10589-017-9954-1

25. K. Bui, F. Park, Y. Lou, J. Xin. A weighted difference of anisotropic and isotropic total variation for relaxed Mumford-Shah color and multiphase image segmentation, *SIAM J. Imaging Sci.*, **14** (2021), 1078–1113. https://doi.org/10.1137/20M1337041

26. Y. Lou, T. Zeng, S, Osher, J. Xin, A weighted difference of anisotropic and isotropic total variation model for image processing, *SIAM J. Imaging Sci.*, **8** (2015), 1798–1823. https://doi.org/10.1137/14098435X

27. Z. Li, Y. Lou, T. Zeng, Variational multiplicative noise removal by DC programming, *J. Sci. Comput.*, **68** (2016), 1200–1216. https://doi.org/10.1007/s10915-016-0175-z

28. Y. Sun, H. Chen, J. Tao, L. Lei, Computed tomography image reconstruction from few views via Log-norm total variation minimization, *Digit. Signal Process.*, **88** (2019), 172–181. https://doi.org/10.1016/j.dsp.2019.02.009

29. B. Zhang, G. Zhu, Z. Zhu, A TV-log nonconvex approach for image deblurring with impulsive noise, *Signal Process.*, **174** (2020), 107631. https://doi.org/10.1016/j.sigpro.2020.107631

30. H. L. Thi, T. P. Dinh, Open issues and recent advances in DC programming and DCA, *J. Global Optim.*, 2023. https://doi.org/10.1007/s10898-023-01272-1

31. T. Wu, Y. Zhao, Z. Mao, L. Shi, Z. Li, T. Zeng, Image segmentation via Fischer-Burmeister total variation and thresholding, *Adv. Appl. Math. Mech.*, **14** (2022), 960–988. https://doi.org/10.4208/aamm.OA-2021-0126

32. M. Lustig, D. Donoho, J. Pauly, Sparse MRI: The application of compressed sensing for rapid MR imaging, *Magn. Reson. Med.*, **58** (2007), 1182–1195. https://doi.org/10.1002/mrm.21391

33. Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli, Image quality assessment: from error visibility to structural similarity, *IEEE T. Image Process.*, **13** (2004), 600–612. https://doi.org/10.1109/TIP.2003.819861

34. K. Zhang, W. Zuo, Y. Chen, D. Meng, L. Zhang, Beyond a gaussian denoiser: Residual learning of deep CNN for image denoising, *IEEE T. Image Process.*, **26** (2017), 3142–3155. https://doi.org/10.1109/TIP.2017.2662206

35. L. Huo, W. Chen, H. Ge, M. K. Ng, Stable image reconstruction using transformed total variation minimization, *SIAM J. Imaging Sci.*, **15** (2022), 1104–1139. https://doi.org/10.1137/120868281

36. Y. Liu, Z. Zhan, J. Cai, D. Guo, Z. Chen, X. Qu, Projected iterative soft-thresholding algorithm for tight frames in compressed sensing magnetic resonance imaging, *IEEE Trans. Med. Imaging*, **35** (2016), 2130–2140. https://doi.org/10.1109/TMI.2016.2550080

37. W. Wang, D. Cao, X. Li, N. Cao, Compressively sampled magnetic resonance imaging reconstruction based on split Bregman iteration with general non-uniform threshold shrinkage, *Magn. Reson. Imaging*, **85** (2022), 297–307. https://doi.org/10.1016/j.mri.2021.10.015