J. Appl. Numer. Optim. 7 (2025), No. 1, pp. 41-54
Available online at http://jano.biemdas.com
https://doi.org/10.23952/jano.7.2025.1.03

DC PROGRAMMING AND ALGORITHM FOR NONCONVEX LOG TOTAL VARIATION IMAGE RECONSTRUCTION

BENXIN ZHANG[1], LEI XUE[1], SHAOHUA SUN[1], HAIXIA LI[1], ZHIBIN ZHU[2,*]

[1]*School of Electronic Engineering and Automation, Guilin University of Electronic Technology, Guilin, China*
[2]*School of Mathematics and Computing Science, Guilin University of Electronic Technology, Guilin, China*

Abstract. In this paper, a nonconvex variational model with log total variation is proposed to investigate robust image reconstruction from a certain number of measurements. Due to the form of the new model, the difference of convex functions (DC) programming is presented and a DC algorithm (DCA) is adopted. The specific DCA subproblem can be solved by the alternating direction method of multipliers (ADMM). Theoretically, we prove the sequence generated by the DCA converges to a stationary point. The experimental results demonstrate that the new method outperforms the competitive methods when the measurement is limited or the noise level is significant.

Keywords. Alternating direction method of multipliers; Difference of convex functions; Image reconstruction; Nonconvex optimization.

2020 Mathematics Subject Classification. 90C26, 90C30, 65K10.

1. INTRODUCTION

In numerous real-world applications, it is necessary to use a small amount of data to perform the task of reconstructing high-quality images. Moreover, the limitation of experimental conditionals may result in the observation data being contaminated by noise. The famous total variation (TV) regularization given by [1] is now widely applied in the field of imaging because of the excellent ability of edge preserving. As the same as [2] analyzed in depth, the advantage of TV model is that it can reconstruct high-quality images from relatively small number of measurements. The TV model has the advantage of being able to reconstruct high-quality images from a relatively small number of measurements. However, it has the disadvantage of creating staircasing artefacts. In order to obtain high-quality images and further reduce artifacts, various models of TV were investigated recently; see, e.g., [3, 4, 5].

We know that the TV norm includes the L_1 norm and gradient operators. According to statistical factors, the L_1 norm may have biased estimations on lager coefficients. Oracle property [6] would be lost. In order to ensure sparseness, the L_0 norm is chosen appropriately instead of the L_1 norm. But, the minimization of L_0 is NP hard. Then, some penalty functions were proposed, which not only can avoid the restriction, but also take advantage of both the L_1 and L_0, such as SCAD [7], Capped-L_1 [8], transformed-L_1 [9], and $L_1 - L_2$ [10]. These nonconvex relaxations

*Corresponding author.
E-mail address: optimization_zhu@163.com (Z. Zhu).
Received 5 September 2024; Accepted 9 February 2025; Published online 20 March 2025.

are superior to L_1 minimization in empirical results, although different algorithms need to be designed.

Recently, plenty of algorithms for convex optimization were introduced and investigated, including the augmented Lagrange multiplier algorithm [11] and the alternating minimization (AM) algorithm [12]. There are also two benchmark methods including ADMM [13] and primal dual approach [14, 15]. In order to solve the nonconvex models, various efficient algorithms were also suggested, such as the proximal AM method [16] and DCA [17]. In [18], a DCA was created for solving the DC programming. It has been widely used and extensively investigated due to its simplicity and efficiency. In [19, 20], Lou et al proved that L_1-L_2 is closer to L_0 than L_1. Then, the regularization model with anisotropic and isotropic TV was developed for magnetic resonance imaging(MRI) and other applications. In [21], a multiplicative denoising model was introduced, and the primal dual gradient method was used to solve non-convex optimization problems whose objective function can be written as the DC functions. In [22], a DCA with adaptive parameters was discussed for impulsive noise removal.

In this paper, we propose a constrained log total variation nonconvex model for image reconstruction, where the constraint matches the statistical properties of the Gaussian noise, i.e., the norm of residual is L_2. Generally, it is difficult to deal with the nonconvex model. To handle this, the original optimization problem is equivalently transformed into DC programming and DCA is naturally adopted. The main contributions are as follows:

First, a new variational model with constraint is proposed, which is nonconvex in terms of regularization terms. The pixels of the image are restricted to $[0,1]$, and the data fidelity term is the L_2 norm of the residual between the observed data and unknown image. In contrast to the unconstrained regularization model, the proposed model does not require adjustment of the regularization parameters.

Second, the DC programming and algorithm are used to solve the nonconvex variational model. Using the special properties of log functions, the objective function in the new model is reformulated as the difference of two convex functions. Then, the original optimization problem is translated into a DC programming. With the help of indicator functions, the global convergence of the DCA is proved. The analysis provides a new idea for the convergence of the DC programming.

Third, experiments on image reconstruction demonstrate that the new method outperforms other competitors, especially, for the scenario where the amount of measurements is limited or the level of noise is significant.

The rests are as follows. Our new model is demonstrated in Section 2. Section 3 presents DC programming and the DCA with ADMM for the proposed model. The convergence analysis of the DCA is given in Section 4. In Section 5, we demonstrate the effectiveness of the new model. In the last section, Section 6, the conclusion is provided.

2. The Model

Given a linear measurement data $y \in R^m$ and $y = \mathscr{M}X + e$, from the original image $X \in R^{N\times N}$, $\mathscr{M} : R^{N\times N} \to R^m$ is the linear operator and $e \in R^m$ is the noise, which is bounded by $\|e\| \leq \tau$ with $\tau \geq 0$. Under the framework of variational approaches, the reconstruction of original image

X is modeled as:

$$\begin{aligned} &\min_{X\in R^{N\times N}} \|X\|_{TV}, \\ &s.t.\ \|\mathscr{M}X-y\| \le \tau, \end{aligned} \tag{2.1}$$

where $\|\cdot\|_{TV}$ is the TV norm and can be categorized as anisotropic:

$$\|\nabla X\|_1 = \|\nabla_x X\|_1 + \|\nabla_y X\|\|_1,$$

and isotropic:

$$\|\nabla X\|_{2,1} = \|\sqrt{|\nabla_x X|^2 + |\nabla_y X|^2}\|_1,$$

where ∇ is the gradient operator for the discrete image, the horizontal and vertical derivative operators are represented by ∇_x and ∇_y, and $\|\cdot\|_1$ is the L_1 norm. So, the TV norm is made up of the gradient operator and the L_1 norm. From the viewpoint of statistics, it would lose the oracle property [6]. Then, some correction procedures and nonconvex approaches are introduced. Inspired by $L_1 - L_2$ minimization [10] in compression sensing problems, Lou et al. [20] gave the famous $L_1 - \gamma L_2$ regularization term by coupling of the anisotropic and isotropic TV for image processing applications:

$$\min_{X\in R^{N\times N}} \|\nabla X\|_1 - \gamma\|\nabla X\|_{2,1} + \frac{\alpha}{2}\|\mathscr{M}X-y\|_2^2,$$

where $0 < \gamma \le 1$ is a parameter and $\alpha > 0$ is a regularization parameter. The authors proposed a DCA for the constrained formulations, and proved its convergence.

In [23], enhanced total variation regularization was discussed as:

$$\begin{aligned} &\min_{X\in R^{N\times N}} \|\nabla X\|_1 - \frac{\gamma}{2}\|\nabla X\|_{2,1}^2, \\ &s.t.\ \|\mathscr{M}X-y\| \le \tau. \end{aligned}$$

The stable reconstruction guarantees were also established for the noisy subsampled measurements with non-adaptive and variable-density sampling. Transformed L1 (TL1) regularization [9] was demonstrated to have comparable signal recovery capability with $L_1 - L_2$ regularization, regardless of whether the measurement matrix satisfies the restricted isometry property. In the spirit of the TL1 method, Huo et al. introduced a transformed total variation (TTV) minimization model [24] to investigate robust image recovery from a certain number of noisy measurements

$$\min_{X\in R^{N\times N}} \frac{(\gamma+1)\|\nabla X\|_1}{\gamma+\|\nabla X\|_1} + \frac{\alpha}{2}\|\mathscr{M}X-y\|_2^2,$$

and the numerical results of image reconstruction illustrated the efficiency of their TTV minimization model.

In this study, the new nonconvex variational model with log penalty is given as

$$\begin{aligned} &\min_{0\le X\le 1} \frac{1}{s}\mathrm{Log}(1+s\|\nabla X\|_1), \\ &s.t.\ \|\mathscr{M}X-y\| \le \tau, \end{aligned} \tag{2.2}$$

where $s > 0$ is the scale parameter and

$$\frac{1}{s}\mathrm{Log}(1+s\|\nabla X\|_1) = \frac{1}{s}\mathrm{Log}(1+s\|\nabla_x X\|_1) + \frac{1}{s}\mathrm{Log}(1+s\|\nabla_y X\|_1).$$

Different from the above models, the pixel values of images in our model are scaled into $[0,1]$, which is more suitable for real applications. Furthermore, we observe that the concave log function is steeper and more closely approximates the L_0 norm if s is small. It promotes sparsity [25] more effectively than the L_1 norm and makes the unbiased estimator.

3. DC Programming and DCA

DC programming and algorithms are powerful tools for nonconvex optimization [17, 26]. We adopt DC technique to solve the proposed model and present the iterative DCA.

First, we split (2.2) as follows:

$$\begin{aligned} &\min_{0\le X\le 1} \Gamma(X) = \Phi(X) - \Psi(X), \\ &s.t.\ \|\mathscr{M}X - y\| \le \tau, \end{aligned} \tag{3.1}$$

where

$$\begin{cases} \Phi(X) = \|\nabla_x X\|_1 + \|\nabla_y X|\|_1, \\ \Psi(X) = \|\nabla_x X\|_1 + \|\nabla_y X|\|_1 - \frac{1}{s}(\mathrm{Log}(1+s\|\nabla_x X\|_1) + \mathrm{Log}(1+s\|\nabla_y X\|_1)). \end{cases}$$

By [27], $\Psi(X)$ is continuously differentiable and convex. Hence, (3.1) is a DC programming problem.

In general, DCA optimizes problem (3.1) by linearizing $\Psi(X)$ as follows:

$$\begin{aligned} X_{k+1} = \arg\min_{0\le X\le 1} \{&\Phi(X) - (\Psi(X_k) + \langle q_x^k, \nabla_x X\rangle) + \langle q_y^k, \nabla_y X\rangle)) + \frac{c}{2}\|X - X_k\|^2, \\ &s.t.\ \|\mathscr{M}X - y\| \le \tau\}, \end{aligned} \tag{3.2}$$

where

$$q_x^k = \frac{s\nabla_x X_k}{1 + \mathrm{sign}(\nabla_y X_k) s\nabla_x X_k}$$

and

$$q_y^k = \frac{s\nabla_y X_k}{1 + \mathrm{sign}(\nabla_y X_k) s\nabla_y X_k},$$

and sign as the signum function. Moreover, the solution X_{k+1} of (3.2) is unique because the objective function is strongly convex. Clearly, convex DCA subproblem (3.2) is the following optimization problem:

$$\begin{aligned} X_{k+1} = \arg\min_{0\le X\le 1} &\|\nabla_x X\|_1 + \|\nabla_y X|\|_1 - (\langle q_x^k, \nabla_x X\rangle + \langle q_y^k, \nabla_y X\rangle) + \frac{c}{2}\|X - X_k\|^2, \\ &s.t.\ \|\mathscr{M}X - y\| \le \tau\}, \end{aligned} \tag{3.3}$$

In order to use ADMM, (3.3) is transformed into the following problem by introducing free variables d_x, d_y, z, V:

$$\begin{aligned} &\min \|d_x\|_1 + \|d_y|\|_1 - (\langle q_x^k, d_x\rangle + \langle q_y^k, d_y\rangle), \\ &s.t.\ \nabla_x X = d_x, \nabla_y X = d_y, \\ &X = V, V \in \mathscr{B}(0,1) := \{0 \le V \le 1\}, \\ &z - (\mathscr{M}X - y) = 0, \\ &z \in \mathscr{Q}(0,\tau) := \{\|z\| \le \tau\}. \end{aligned}$$

The augmented Lagrangian function of (3.3) is

$$\begin{aligned}
&\mathscr{L}(X,d_x,d_y,z,b_x,b_y)\\
&=\|d_x\|_1+\|d_y\|\|_1-\langle q_x^k,d_x\rangle-\langle q_y^k,d_y\rangle+\frac{\mu}{2}\|z-(\mathscr{M}X-y)-\rho\|^2\\
&\quad+\frac{\lambda}{2}\|d_x-\nabla_x X-b_x\|^2+\frac{\lambda}{2}\|d_y-\nabla_y X-b_y\|^2,\\
&\quad+\frac{\beta}{2}\|V-X-\nu\|^2+\frac{c}{2}\|X-X_k\|^2.
\end{aligned}\tag{3.4}$$

where $\mu,\beta,\lambda>0$ are penalty parameters, and ρ, ν, b_x, and b_y are Lagrange multipliers. Next, alternating minimization for the above Lagrangian consists of solving X,d_x,d_y,z,V subproblems, and updating the dual variables. Below we elaborate on solving each subproblem in (3.4) with j-th iteration.

The solution of subproblem X can be expressed as:

$$X^{j+1}=\arg\min\mathscr{L}(X,d_x^j,d_y^j,b_x^j,b_y^j).\tag{3.5}$$

By the optimality condition of (3.5), we have

$$\begin{aligned}
&(\mu\mathscr{M}^T\mathscr{M}+\lambda\nabla^T\nabla+c+\beta)X^{j+1}\\
&=\lambda\nabla_x^T(d_x^j-b_x^j)+\lambda\nabla_y^T(d_y^j-b_y^j)+\mu\mathscr{M}^T(y-z^j-\rho^j)+cX_k+\beta(V^j-\nu^j).
\end{aligned}\tag{3.6}$$

In the compressive sensing undersampling MRI reconstruction problem [28], $\mathscr{M}$ is composed of the Fourier transform F and sampling operator R, i.e., $\mathscr{M}=RF$. When the periodic boundary condition is applied, $\mu\mathscr{M}^T\mathscr{M}+\lambda\nabla_x^T\nabla_x+\lambda\nabla_y^T\nabla_y$ can be diagonalized by the fast Fourier transform (FFT). Hence, (3.6) is efficiently solved by FFT: $X^{j+1}=$

$$\mathscr{F}^{-1}\Big(\frac{\mathscr{F}(\lambda\nabla_x^T(d_x^j-b_x^j))+\mathscr{F}(\lambda\nabla_y^T(d_x^j-b_x^j))+\mu\mathscr{F}(\mathscr{M}^T(y-z^j-\rho^j))+c\mathscr{F}(X_k)+\beta\mathscr{F}(V^j-\nu^j)}{\mu\mathscr{F}(\mathscr{M})^*\odot\mathscr{F}(\mathscr{M})+\lambda\mathscr{F}(\nabla)^*\odot\mathscr{F}(\nabla)+\mathscr{F}(c+\beta)}\Big),$$

where $\mathscr{F}^{-1}$ and $\mathscr{F}$ represent inverse FFT and FFT, $*$ denotes conjugation, and $\odot$ refers to the elementwise multiplication.

The optimal solutions of d_x,d_y-subproblems are obtained by the shrinkage operator:

$$d_x^{j+1}=\text{shrink}(\nabla_x X^{j+1}+b_x^j+q_x^k/\lambda,1/\lambda)\tag{3.7}$$

and

$$d_y^{j+1}=\text{shrink}(\nabla_y X^{j+1}+b_y^j+q_y^k/\lambda,1/\lambda),\tag{3.8}$$

where

$$\text{shrink}(t_1,t_2)=\text{sign}\max\{|t_1|-t_2,0\}.$$

For the z-subproblem, the solution can be obtained by the projection operator

$$z^{j+1}=\mathscr{P}_{\mathscr{B}(0,\tau)}(\mathscr{M}X^{j+1}-y+\rho^j).\tag{3.9}$$

For the V-subproblem, the solution can also be obtained by the projection operator

$$V^{j+1}=\mathscr{P}_{\mathscr{B}(0,1)}(X^{j+1}+\nu^j).\tag{3.10}$$

The multipliers of b_x, b_y ρ, and ν are expressed as:

$$b_x^{j+1}=b_x^j+D_x u^{j+1}-d_x^{j+1},\tag{3.11}$$

$$b_y^{j+1}=b_y^j+D_y u^{j+1}-d_y^{j+1},\tag{3.12}$$

$$\rho^{j+1} = \rho^j + \mathscr{M}X^{j+1} - y - z^{j+1}, \tag{3.13}$$

and

$$\nu^{j+1} = \nu^j + X^{j+1} - V^{j+1}. \tag{3.14}$$

In fact, the above schemes for solving subproblem (3.3) is a straightforward application of ADMM to X and (d_x, d_y, z, V). Then, by [29, Theorem 3.2], ADMM for two blocks of variables converges.

Finally, we give our DCA for solving (2.2) in Algorithm 1.

Algorithm 1 DCA for solving (2.2).

Set $X_0 = d_x^0 = d_y^0 = z^0 = V^0 = \rho^0 = 0$, MaxDCA, j_{max},,
For $k = 0, 1, 2, \cdots$, MaxDCA, $b_x^0 = b_y^0 = \rho^0 = \nu^0 = 0$,
For $j = 0, 1, 2, \cdots, j_{max}$
Compute X^{j+1} by (3.6),
Compute d_x^{j+1} by (3.7),
Compute d_y^{j+1} by (3.8),
Compute z^{j+1} by (3.9),
Compute V^{j+1} by (3.10),
Update b_x^{j+1} by (3.11),
Update b_y^{j+1} by (3.12),
Update ρ^{j+1} by (3.13).
Update ν^{j+1} by (3.14).
End For
$X_{k+1} = X^{j+1}$,
$q_{k+1} = (q_x^{k+1}, q_y^{k+1}) = (\frac{s\nabla_x X_{k+1}}{1+\text{sign}(\nabla_y X_{k+1})s\nabla_x X_{k+1}}, \frac{s\nabla_y X_{k+1}}{1+\text{sign}(\nabla_y X_{k+1})s\nabla_y X_{k+1}})$. End For

4. Convergence Analysis

In this section, we show that the proposed DCA converges to a stationary point by using the toll of variational analysis. First, we introduc the indicator function

$$\delta_{\mathscr{B}(0,1)}(X) = \begin{cases} 0, & 0 \le X \le 1, \\ +\infty, & \text{otherwise} \end{cases}$$

and

$$\delta_{\mathscr{Q}(0,\tau)}(X) = \begin{cases} 0, & \|\mathscr{M}X - y\| \le \tau, \\ +\infty, & \text{otherwise}, \end{cases}$$

and we rewrite (3.1) as an unconstrained problem

$$\min_X \Gamma(X) = \Phi(X) - \Psi(X) + \delta_{\mathscr{Q}(0,\tau)}(X) + \delta_{\mathscr{B}(0,1)}(X). \tag{4.1}$$

Let

$$g(X) = \Phi(X) + \delta_{\mathscr{Q}(0,\tau)}(X) + \delta_{\mathscr{B}(0,1)}(X),\ h(X) = \Psi(X).$$

Now, we present that the object function in our method is decreased.

Lemma 4.1. *Let $\{X_k\}$ be the sequence generated by Algorithm 1. Then*

$$\Gamma(X_k) - \Gamma(X_{k+1}) \geq c\|X_{k+1} - X_k\|^2.$$

Proof. By the convexity of $g(X)$, one has

$$g(X_k) - g(X_{k+1}) \geq \langle \partial g(X_{k+1}), X_k - X_{k+1} \rangle. \tag{4.2}$$

Note that

$$h(X_{k+1}) - h(X_k) \geq \langle \partial h(X_k), X_{k+1} - X_k \rangle. \tag{4.3}$$

Moreover, by the first-order optimality condition of (3.3), one sees that there exists $p_{k+1} \in \partial(\|\nabla X\|_1 + \delta_{\mathscr{Q}(0,\tau)}(X) + \delta_{\mathscr{B}(0,1)}(X))$ such that

$$0 = p_{k+1} - \nabla^T q_k + c(X_{k+1} - X_k) \tag{4.4}$$

Combining (4.2), (4.3) and (4.4), one arrives at

$$\Gamma(u_k) - \Gamma(u_{k+1}) \geq c\|X_{k+1} - X_k\|_2^2.$$

The assertion is proven. □

Theorem 4.1. *Let $\{X_k\}$ be the sequence generated by Algorithm 1. Then it converges to a critical point of (4.1).*

Proof. $\{\Gamma(X_k)\}$ is bounded below ($\Gamma(X) \geq 0$) and decreased. Obviously,

$$\sum_{k=0}^{\infty} \|X_{k+1} - X_k\|^2 \leq \Gamma(X_0)/c.$$

Then, $\lim_{k\to\infty}(X_{k+1} - X_k) = 0$. In view of $0 \leq X_k \leq 1$, one sees that $\{X_k\}$ is bounded. Assume that X^* is any accumulation point of X_k. Then there exists a subsequence X_{n_k} such that $\lim_{k\to\infty} X_{n_k} = X^*$. Since X_{n_k} is generated by Algorithm 1, all of its entries are bounded by $\mathscr{P}_{\mathscr{Q}(0,\tau)}$; otherwise, the objective function would be at $+\infty$. Hence, $\delta_{\mathscr{Q}(0,\tau)}(X_{n_k}) = 0$ and also $\delta_{\mathscr{Q}(0,\tau)}(X_{n_k+1}) = 0$. It follows that

$$\lim_{n_k\to\infty} \delta_{\mathscr{Q}(0,\tau)}(X_{n_k}) = \delta_{\mathscr{Q}(0,\tau)}(X^*),$$

and

$$\lim_{n_k\to\infty} \delta_{\mathscr{B}(0,1)}(X_{n_k}) = \delta_{\mathscr{B}(0,1)}(X^*).$$

Since $\Psi(X)$ is continuously differentiable and ∇ is continuous (linear operator), then

$$\lim_{n_k\to\infty} \nabla^T q_{n_k} = \nabla^T q^* \in \partial\Psi(X^*).$$

By (4.4), one sees that

$$0 \in \partial(\|\nabla X_{n_k+1}\|_1 + \delta_{\mathscr{Q}(0,\tau)}(X_{n_k+1}) + \delta_{\mathscr{B}(0,1)}(X_{n_k+1})) - \nabla^T q_{n_k} + c(X_{n_k+1} - X_{n_k}).$$

It further follows the subgradient that

$$\begin{aligned}\|\nabla X_k\|_1 + \delta_{\mathscr{Q}(0,\tau)}(X_k) + \delta_{\mathscr{B}(0,1)}(X_k) \geq{}& \|\nabla X_{n_k+1}\|_1 + \delta_{\mathscr{Q}(0,\tau)}(X_{n_k+1}) + \delta_{\mathscr{B}(0,1)}(X_{n_k+1}) \\ &+ \langle \nabla^T q_{n_k} - c(X_{n_k+1} - X_{n_k}), X_k - X_{n_k+1} \rangle.\end{aligned}$$

In view of the continuity, one obtains

$$\begin{aligned}
&\|\nabla X_k\|_1 + \delta_{\mathscr{Q}(0,\tau)}(X_k) + \delta_{\mathscr{B}(0,1)}(X_k) \\
&\geq \lim_{n_k\to\infty} (\|\nabla X_{n_k+1}\|_1 + \delta_{\mathscr{Q}(0,\tau)}(X_{n_k+1}) + \delta_{\mathscr{B}(0,1)}(X_{n_k+1}) \\
&\quad + \langle \nabla^T q_{n_k} - c(X_{n_k+1} - X_{n_k}), X_k - X_{n_k+1}\rangle) \\
&= \|\nabla X^*\|_1 + \delta_{\mathscr{Q}(0,\tau)}(X^*) + \delta_{\mathscr{B}(0,1)}(X^*) + \langle \nabla^T q^*, X_k - X^*\rangle.
\end{aligned}$$

From the convexity, one has

$$\begin{aligned}
\nabla^T q^* &\in \partial(\|\nabla X^*\|_1 + \delta_{\mathscr{Q}(0,\tau)}(X^*) + \delta_{\mathscr{B}(0,1)}(X^*)) \\
&= \partial\|\nabla X^*\|_1 + \partial\delta_{\mathscr{Q}(0,\tau)}(X^*) + \partial\delta_{\mathscr{B}(0,1)}(X^*).
\end{aligned}$$

Thus

$$\begin{aligned}
0 &\in \partial\|\nabla X^*\|_1 + \partial\delta_{\mathscr{Q}(0,\tau)}(X^*) + \partial\delta_{\mathscr{B}(0,1)}(X^*) - \nabla^T q^* \\
&\subseteq \partial\Phi(X^*) + \partial\delta_{\mathscr{Q}(0,\tau)}(X^*) + \partial\delta_{\mathscr{B}(0,1)}(X^*) - \partial\Psi(X^*).
\end{aligned}$$

□

5. EXPERIMENTS

In this section, experimental results are reported to validate the effectiveness of model (2.2) and Algorithm 1. We use the TV (2.1), $L_1 - 0.5L_2$ [20], TTV [24], and ETV [23] models for comparison. TV model (2.1) is solved by ADMM. $L_1 - 0.5L_2$, TTV , and ETV models are solved by the DCA and the corresponding subproblem is solved by ADMM. The suggested parameter settings in compared methods are used according to the related references for fair comparison. In Figure 1, we show the tested images including the Shepp-Logan image, shape and Texmos3 (synthetic piecewise-constant images), House and Lena (natural images), and Brain and Mr030 (medical images).

Two radial lines and random sampling strategies are considered. We conduct experiments on a desktop with 8 GB RAM and i5-4690 Processor. The peak signal-to-noise ratio (PSNR) and the structure similarity (SSIM) [30] are used to evaluate the restored quality. For algorithmic parameters, we take for $s = 200$, $c = 0.001$, $\beta = 0.00001$, $\mu = 10$, and $\lambda = 100$. In most cases, it can achieve good performance. The starting points for all methods are set as zeros. We use the stopping criterion $\|X_k - X_{k-1}\| \leq 10^{-10}$ for running the DCA, and the maximum outer and inner iteration numbers are 15 and 200.

In the field of image reconstruction, one of the standard images is the Shepp-Logan phantom. The experiments on this standard test image are split into two parts. The first focuses on reconstructing the Shepp-Logan (256×256) from measurements with no noise. First,7 and 8 radial lines are used for sampling and the corresponding sampling rates are 3.03% and 3.98%. Moreover, we also take random measurements with 1.90% and 1.50%. As demonstrated in Figure 2, the proposed model can produce accurate reconstruction for the four sampling settings, and the qualities are much better than the other models when the amounts of samples are limited.

Table 1 presents the PSNR and SSIM values. The superiority of the proposed model is evident when the measurements are limited. Our model obtains the highest SSIM value (0.9197) when the sampling rate is 1.5%. We also note that though the SSIM values are 1 for the proposed model and ETV in three sampling settings, the PSNR values of the new model are higher than

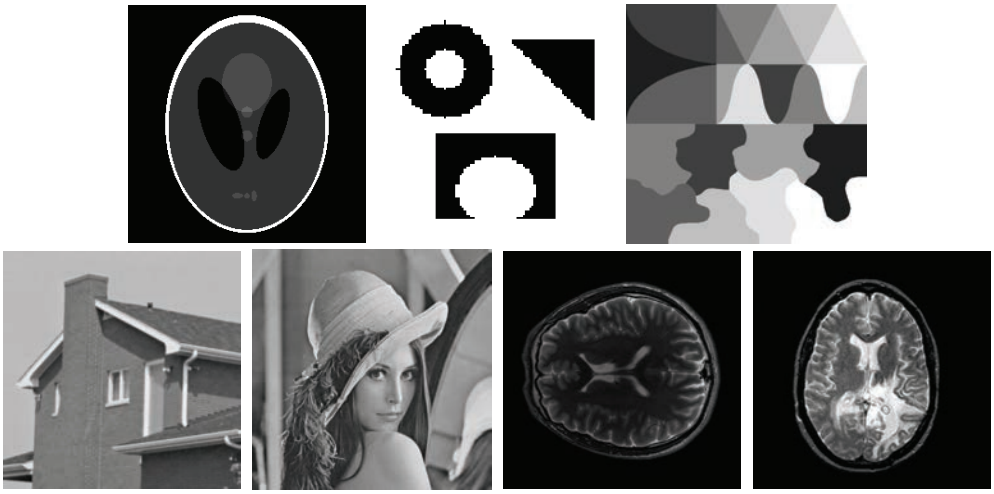

FIGURE 1. Test images.

TABLE 1. SSIM and PSNR of the five models on Shepp-Logan image.

Mask	TV		$L_1-0.5L_2$		TTV		ETV		LogTV	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
7 lines	18.4424	0.4528	21.2343	0.6616	23.3042	0.8461	66.4012	1.0000	84.0535	1.0000
8 lines	24.2292	0.7150	177.2076	1.000	57.2630	0.9396	79.1119	1.0000	94.3026	1.0000
1.50%	22.5254	0.8382	23.0122	0.8447	23.0737	0.8377	23.7606	0.8965	24.4769	0.9197
1.90%	27.9968	0.9439	39.0418	0.9827	50.7258	0.9982	64.6380	1.0000	89.0146	1.0000

TABLE 2. PSNR and SSIM values of the five models on Shepp-Logan image with levels of noise σ=0.05, 0.1.

mask	σ	TV		$L_1-0.5L_2$		TTV		ETV		LogTV	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
15 lines	0.05	25.1430	0.6199	27.3036	0.5727	29.3038	0.6826	32.9900	0.9352	31.0499	0.9547
	0.1	20.9155	0.4220	22.6635	0.3673	23.0592	0.4547	29.0121	0.7546	26.8304	0.9213
1.90%	0.05	22.8989	0.7358	23.5717	0.7315	24.6205	0.8330	21.3315	0.8743	22.4498	0.9039
	0.1	20.8328	0.5488	21.3376	0.5690	21.5555	0.6960	20.9260	0.8237	22.1929	0.8931

TABLE 3. PSNR and SSIM of the five models on Shape and Texmos3 images.

Image	Mask	TV		$L_1-0.5L_2$		TTV		ETV		LogTV	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Shape	3 lines	11.3273	0.5171	13.1152	0.5466	20.6603	0.8127	28.1186	0.9662	33.1541	0.9877
Texmos3	6 lines	25.8484	0.9586	28.3377	0.9629	35.9768	0.9819	42.4298	0.9992	65.2897	1.0000

the ETV model. It is well illustrated that our new model is stable with respect to the limited measurements for reconstructing the Shepp-Logan image.

The second illustrates the robustness of the proposed model with respect to noise. We take measurements along 15 lines (6.44%) and use a 1.9% sampling mask. Gaussian noises are perturbed for the Fourier measurements with standard derivations σ of 0.05 and 0.1. The PSNR and SSIM values are displayed in Table 2. In terms of tge SSIM value, the new method has the highest reconstruction quality. Also, when the level of noise increases, the superiorities are more apparent. Hence, the results demonstrate that the proposed approach is robust for noise.

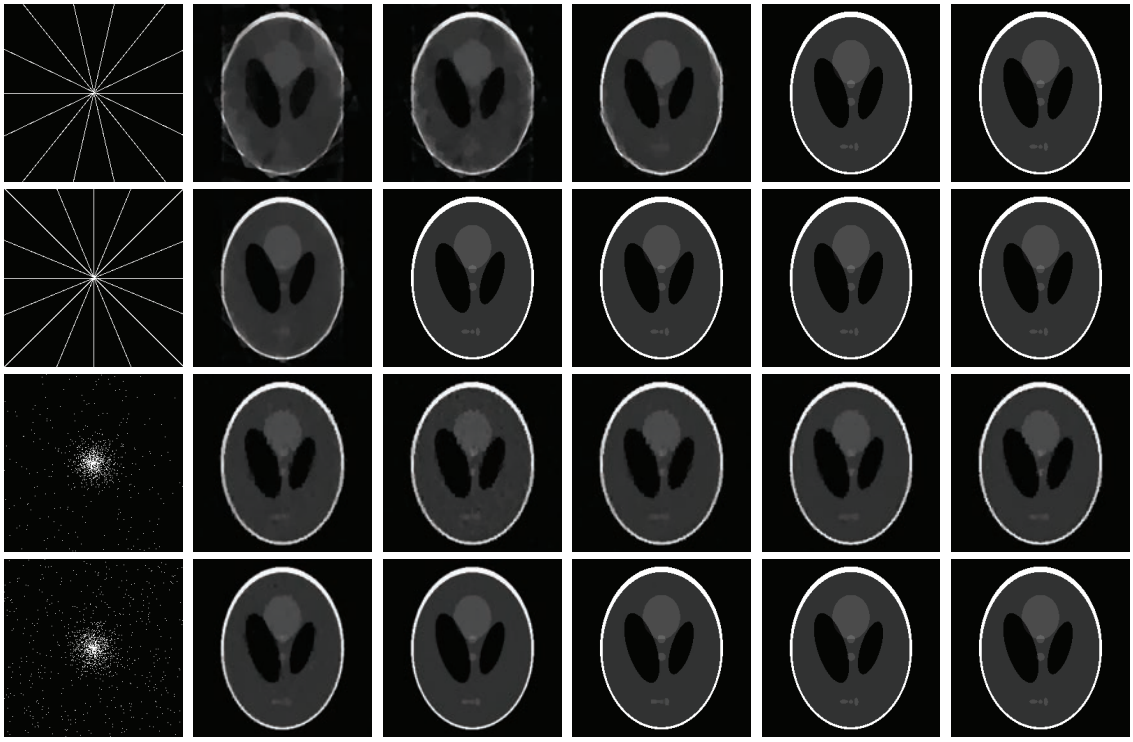

FIGURE 2. Shepp-Logan: comparison of TV, $L_1 - 0.5L_2$, TTV, ETV and LogTV models with radial lines (first row: 7 lines, second row: 8 lines) and random (third row: 1.5% sampling rates, fourth row: 1.9% sampling rates) masks.

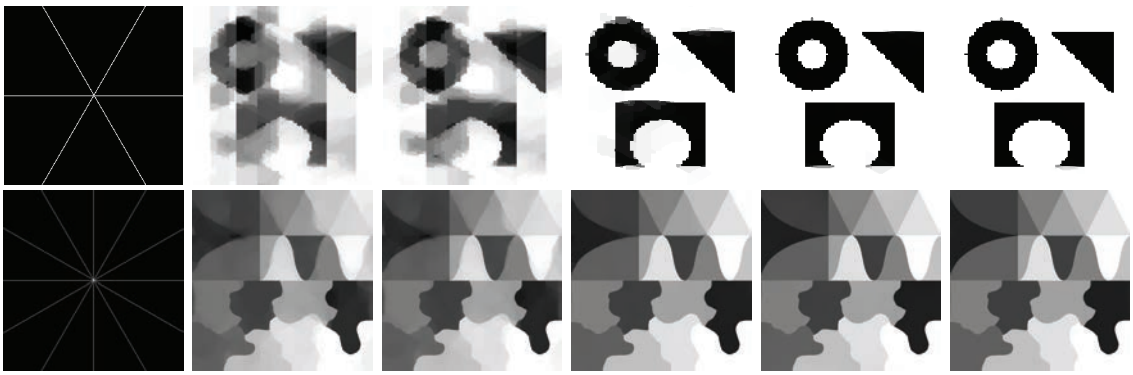

FIGURE 3. Shape and Texmos3: comparison of TV, $L_1 - 0.5L_2$, TTV, ETV and LogTV models with radial lines masks (first row: 3 lines, second row: 6 lines).

Next, we focus on the radial line sampling and verify the superiority by conducting experiments on two synthetic images: Shape and Texmos3. When we have a limited number of measurements, for example, 3 or 6 lines, our model can also generate good reconstructions. In Figure 3, we show the reconstruction results. In Table 3, the SSIM and PSNR values are listed. From Table 3 and Figure 3, one sees that the results of the proposed model and algorithm are significantly better than the other methods.

TABLE 4. PSNR and SSIM of the five models on House and Lena images.

Image	Mask	TV		$L_1 - 0.5L_2$		TTV		ETV		LogTV	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
House	10%	29.5784	0.8310	30.7095	0.8481	30.6767	0.8485	30.7715	0.8493	30.9273	0.8540
	15%	30.5248	0.8428	30.7095	0.8481	32.0679	0.8485	32.2181	0.8667	32.7832	0.8750
Lena	10%	27.7920	0.8039	29.0749	0.8184	28.7771	0.8229	28.8094	0.8207	29.2737	0.8254
	15%	28.8077	0.8228	30.3036	0.8452	30.1061	0.8498	30.1832	0.8477	30.9159	0.8567

FIGURE 4. House and Lena: comparison of TV, $L_1 - 0.5L_2$, TTV, ETV and LogTV models with random masks (first row: 10% sampling rates, second row: 15% sampling rates).

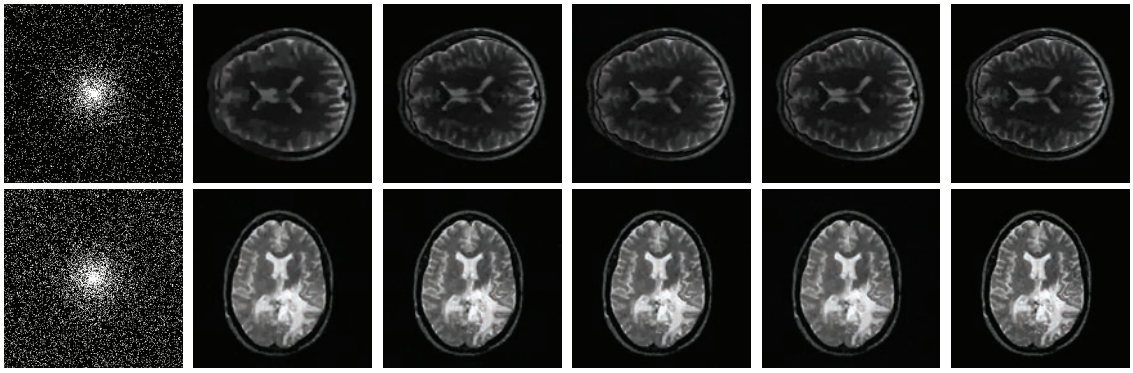

FIGURE 5. Brain and Mr030: comparison of TV, $L_1 - 0.5L_2$, TTV, ETV and LogTV models with random masks (first row: 10% sampling rates, second row: 15% sampling rates).

TABLE 5. PSNR and SSIM of the five models on Brain and Mr030 images.

Image	Mask	TV		$L_1 - 0.5L_2$		TTV		ETV		LogTV	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Brain	10%	26.4954	0.8086	28.3430	0.8627	29.2273	0.8001	28.6943	0.8643	30.0639	0.8958
	15%	27.0739	0.8279	29.5160	0.8889	30.5752	0.8217	30.1018	0.8932	30.9309	0.9231
Mr030	10%	27.5838	0.7689	29.0406	0.7849	28.8255	0.8747	28.7354	0.8208	30.0639	0.9128
	15%	28.9121	0.8170	30.9751	0.8218	29.8584	0.9234	30.9548	0.8342	32.9545	0.9469

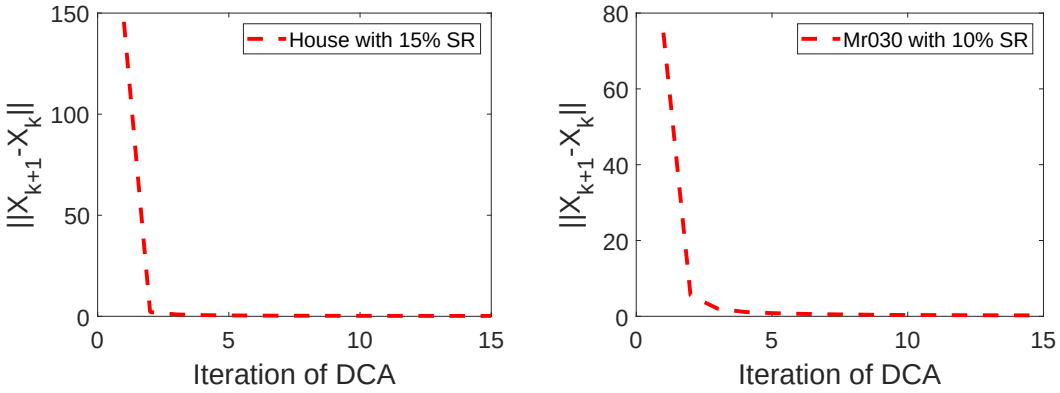

FIGURE 6. Curves of error versus outer iteration numbers.

We further validate the superiority of the new method by conducting experiments on two natural images: House and Lena. In Figure 4, the vision results of reconstruction are displayed. We also report the PSNR and SSIM values for each reconstruction from the random sampling rates of 10% and 15% in Table 4. Because the edges of natural images are more complicated, it is worth noting that the performance of new method for the natural images may be not effective enough as its performance for the synthetic images. Nonetheless, the proposed model and algorithm get the highest PSNR and SSIM values in all methods.

We also apply the new approach to two medical images: Brain and Mr030. We use the random sampling rates of 10% and 15% for the test. In Figure 5 and Table 5, the reconstructed images and results are displayed. Compared with other methods, it is obvious that the new method produces better reconstructions. We also observe that the superiority of the new method is more apparent when the sampling rates are relatively low. Hence, our proposed model is preferred when the measurements are limited.

At last, we verify the convergence of convergence of ADMM (inner iteration) and DCA (outer interation). We choose House image with sampling rates 15% and Mr030 image with sampling rates 10% as examples. Figure 6 presents the curves of convergence about the proposed DCA, illustrating the error between X_k and X_{k+1} over outer iteration numbers. We can see that the curves of error values decrease as the outer iteration numbers increase and become flat after a period of time. Hence, the proposed DCA is stable and has a good convergence property. Figure 7 shows the error between X^j and X^{j+1} of the inner iteration sequence generated by ADMM when the numbers of outer iteration k are 1 and 15 with the number of inner iteration j^{max} is 200. It can be seen that when the number of inner iteration is set to 200, ADMM has obtained a good solution of DCA subproblem (3.3), which can be considered as the optimal solution because the error curve has become flat and the error value is very small. Based on the guidance provided in Figures 6 and 7, this is the reason that why we set the maximum outer and inner iterations numbers are 15 and 200, and they can ensure the convergence of ADMM and DCA with theoretically and numerically in our experiments.

6. CONCLUSION

This paper introduces a novel DC programming and algorithm for solving the nonconvex log TV image reconstruction model. The experiments demonstrate that our proposed model and

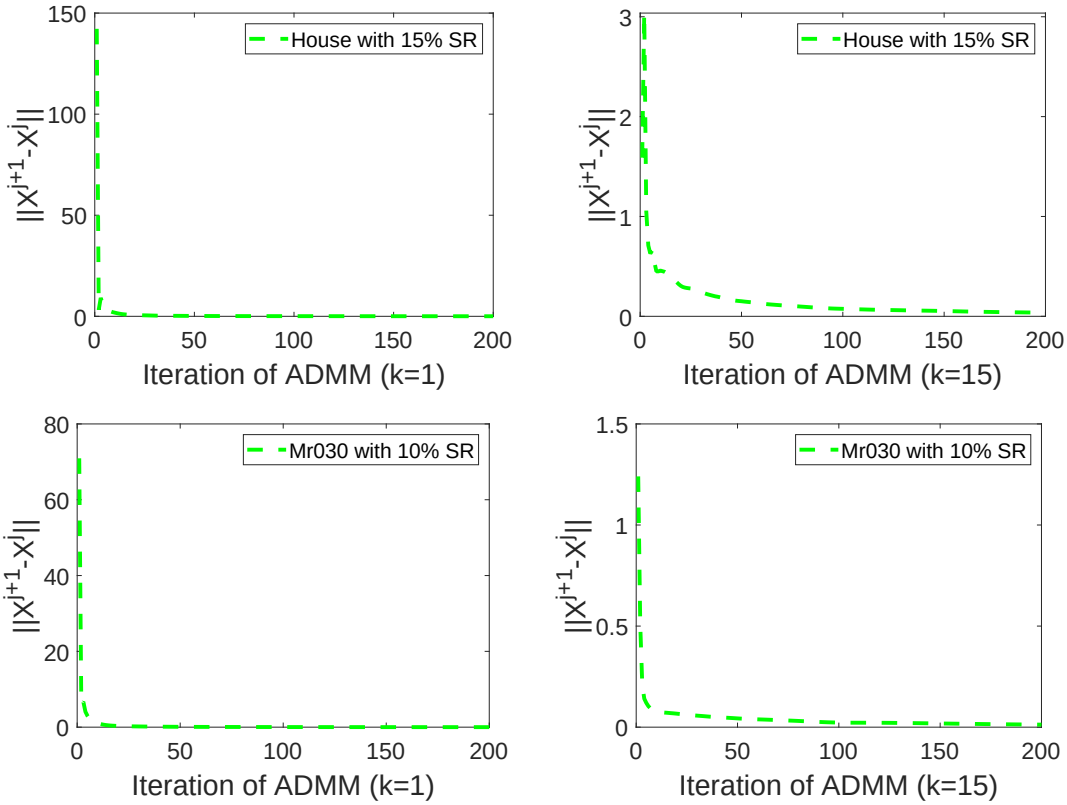

FIGURE 7. Curves of error versus inner iteration numbers.

algorithm are capable of achieving better solutions. In the future, with other data fidelity terms, we may consider some different object functions to deal with noise, for example, impulsive or Poisson noise. The new approach can be applied to higher dimensional images as well. Furthermore, it seems promising to apply the new algorithm to image inpainting or segmentation, phase retrieval, and super-resolution problems.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grant 62171147 and Guangxi Natural Science Foundation under Grant 2024GXNSFAA010512.

REFERENCES

[1] L. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms, Physica D: Nonlinear Phenomena, 60 (1992), 259-268.

[2] D. Needell, R, Ward, Stable image reconstruction using total variation minimization, SIAM Journal on Imaging Sciences, 6 (2013), 1035-1058.

[3] K. Bredies, K. Kunisch, T. Pock, Total generalized variation, SIAM Journal on Imaging Sciences, 3 (2010), 492-526.

[4] L. Condat, Discrete total variation: New definition and minimization, SIAM Journal on Imaging Sciences, 10 (2017), 1258-1290.

[5] Z. Jia , M.K. Ng, W. Wang, Color image restoration by saturation-value total variation, SIAM Journal on Imaging Sciences, 12 (2019), 972-1000.

[6] R. Tibshirani, Regression shrinkage and selection via the lasso, Journal of the Royal Statistical Society: Series B (Methodological), 58 (1996), 267-288.
[7] J. Fan, H. Peng, Nonconcave penalized likelihood with a diverging number of parameters, The Annals of Statistics, 32 (2004), 928-961.
[8] T. Zhang, Analysis of multi-stage convex relaxation for sparse regularization, Journal of Machine Learning Research, 11 (2010), 1081-1107.
[9] S. Zhang, J. Xin, Minimization of transformed L1 penalty: theory, difference of convex function algorithm, and robust application in compressed sensing, Mathematical Programming, 169 (2018), 307-336.
[10] Y. Lou, P. Yin, Q. He, J. Xin, Computing sparse representation in a highly coherent dictionary based on difference of l1 and l2, Journal of Scientific Computing, 64 (2015), 178–196.
[11] M. Fortin, R. Glowinski, On decomposition-coordination methods using an augmented lagrangian, Studies in Mathematics and Its Applications, 15 (1983), 97-146.
[12] Y. Wang, J. Yang , W. Yin, Y. Zhang, A new alternating minimization algorithm for total variation image reconstruction, SIAM Journal on Imaging Sciences, 1 (2008), 248-272.
[13] S, Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, Distributed optimization and statistical learning via the alternating direction method of multipliers, Foundations and Trends in Machine learning, 3 (2011), 1-122.
[14] A. Chambolle, T. Pock, A first-order primal-dual algorithm for convex problems with applications to imaging, Journal of Mathematical Imaging and Vision, 40 (2011), 120-145.
[15] B. Zhang, Z. Zhu, C. Xu, A primal-dual multiplier method for total variation image restoration, Applied Numerical Mathematics, 145 (2019), 145-158.
[16] J. Bolte, S. Sabach, M. Teboulle, Proximal alternating linearized minimization for nonconvex and nonsmooth problems, Mathematical Programming, 146 (2014), 459-494.
[17] T. H. Le, T. Dinh, DC programming and DCA: thirty years of developments, Mathematical Programming, 169 (2018), 5-68.
[18] D. T. Pham, S. E. Bernoussi, Algorithms for solving a class of nonconvex optimization problems, methods of subgradients, North-Holland Math. Stud. vol. 129, pp. 249-271, North-Holland, Amsterdam, 1986.
[19] 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 Journal on Imaging Sciences, 14 (2021), 1078-1113.
[20] Y. Lou, T. Zeng, S. Osher, J. Xin, A weighted difference of anisotropic and isotropic total variation model for image processing, SIAM Journal on Imaging Sciences, 8 (2015), 1798-1823.
[21] Z. Li, Y. Lou, T. Zeng, Variational multiplicative noise removal by DC programming, Journal of Scientific Computing, 68 (2016), 1200-1216.
[22] B. Zhang, G. Zhu, Z. Zhu, A TV-log nonconvex approach for image deblurring with impulsive noise, Signal Processing, 174 (2020), 107631.
[23] C. An, H. N. Wu, X. Yuan, Enhanced total variation minimization for stable image reconstruction, Inverse Problems, 39 (2023), 075005.
[24] L. Huo, W. Chen, H. Ge, M. K. Ng, Stable image reconstruction using transformed total variation minimization, SIAM Journal on Imaging Sciences, 15 (2022), 1104-1139.
[25] E. J. Candes, M. B. Wakin, S. Boyd, Enhancing sparsity by reweighted L_1 minimization, Journal of Fourier Analysis and Applications, 14 (2008), 877-905.
[26] H. A. Le Thi, D. T. Pham, Open issues and recent advances in DC programming and DCA, Journal of Global Optimization, 88 (2024), 533–590.
[27] I. Selesnick, Sparse regularization via convex analysis, IEEE Transactions on Signal Processing, 65 (2017), 4481-4494.
[28] M. Lustig, D. Donoho, J. Pauly, Sparse MRI: The application of compressed sensing for rapid MR imaging, Magnetic Resonance in Medicine, 58 (2007), 1182-1195.
[29] R. H. Chan, M. Tao, X. Yuan, Constrained total variation deblurring models and fast algorithms based on alternating direction method of multipliers, SIAM Journal on Imaging Sciences, 6 (2013), 680-697.
[30] Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli, Image quality assessment: from error visibility to structural similarity, IEEE Transactions on Image Processing, 13 (2004), 600-612.