

# AUGMENTED LAGRANGIAN METHOD, DUAL METHODS, AND SPLIT BREGMAN ITERATION FOR ROF, VECTORIAL TV, AND HIGH ORDER MODELS

CHUNLIN WU \* AND XUE-CHENG TAI \*†

**Abstract.** In image processing, the Rudin-Osher-Fatemi (ROF) model [L. Rudin, S. Osher, and E. Fatemi, *Physica D*, 60(1992), pp. 259–268] based on total variation (TV) minimization has proven to be very useful. A lot of efforts have been devoted to obtain fast numerical schemes and overcome the non-differentiability of the model. Methods considered to be particularly efficient for the ROF model include the dual methods of Chan-Golub-Mulet (CGM) [T.F. Chan, G.H. Golub, and P. Mulet, *SIAM J. Sci. Comput.*, 20(1999), pp. 1964–1977] and Chambolle [A. Chambolle, *J. Math. Imaging Vis.*, 20(2004), pp. 89–97], and splitting and penalty based method [Y. Wang, J. Yang, W. Yin, and Y. Zhang, *SIAM J. Imaging Sciences*, 1(2008), pp. 248–272], as well as split Bregman iteration [T. Goldstein, and S. Osher, *SIAM J. Imaging Sciences*, 2(2009), pp. 323–343]. In this paper, we propose to use augmented Lagrangian method to solve the model. Convergence analysis will be given for the method. In addition, we observe close connections between the method proposed here and some of the existing methods. We show that the augmented Lagrangian method, dual methods, and split Bregman iteration are different iterative procedures to solve the same system. Moreover, the proposed method is extended to vectorial TV and high order models. Using the approach here, we can easily obtain the CGM dual method and split Bregman iteration for vectorial TV and high order models, which, to our best of knowledge, have not been presented in the literature. Numerical examples demonstrate the efficiency and accuracy of our method.

**Key words.** augmented Lagrangian method, dual method, split Bregman iteration, ROF model, total variation

**AMS subject classifications.**

**1. Introduction.** Image restoration such as denoising and deblurring are the most fundamental tasks in image processing. To preserve image edges and features in image regularization is difficult but very desired. Recently, the ROF model [33] has been demonstrated very successful in edge-preserving image restoration. The model immediately attracted much attention and has been extended to high order models [13, 46, 27, 29, 24, 35] and vectorial models for color image restoration [34, 2, 4, 14]; see [15] for an overview.

However, the numerical computation of the ROF model suffers from difficulties related to its nonlinearity and non-differentiability. In [33], the authors proposed a time marching strategy to the associated Euler-Lagrange equation. This method is slow due to the constraint of stability conditions about the time step size. To find fast algorithms has been an active research area so far.

There are several methods that have proven to be particularly efficient for image restoration problems based on the ROF model. One class of approaches is dual methods [12, 9, 11, 48], which are based on dual formulation of the ROF model. The other is based on variable-splitting and equality constrained optimization, e.g., the approach proposed in [39, 40, 42] which uses alternative minimization of the penalized cost functional, and the method in [25] where splitting is applied to the data fidelity term, as well as split Bregman iteration [43, 22]. In this paper, we use a technique related to the augmented Lagrangian method to solve the ROF model. Convergence analysis of the proposed approach will be supplied. In addition, we show that the

---

\*Division of Mathematical Sciences, School of Physical & Mathematical Sciences, Nanyang Technological University, Singapore.

†Department of Mathematics, University of Bergen, Norway.

augmented Lagrangian method, the dual methods, and split Bregman iteration are just different iterative schemes to solve the same system. Some connections between CGM and Chambolle's dual methods have been noticed in [48]. In the context of compressive sensing, the authors in [43] pointed out the equivalence between augmented Lagrangian method and Bregman iteration. Here we show how CGM and Chambolle's dual methods are connected to the augmented Lagrangian method, and verify the equivalence between augmented Lagrangian method and split Bregman iteration. The proposed method is extended to vectorial TV and high order models. Using the extension, we can easily get dual methods and split Bregman iteration for vectorial TV and high order models. To our knowledge, the CGM dual method and split Bregman iteration for vectorial TV and high order models are still missing in the literature.

The paper is organized as follows. In the next section, we give basic notations. In Section 3, we present the ROF model and some existing solvers. Augmented Lagrangian method will be given in Section 4 with some convergence analysis. In Section 5, we show connections between the proposed method and dual methods as well as split Bregman iteration. Our approach and observations are then extended to vectorial TV in Section 6 and high order models in Section 7. Finally, we present some numerical experiments and conclude the paper.

**2. Basic notations.** Without the loss of generality, we represent a gray image as an  $N \times N$  matrix. The Euclidean space  $\mathbb{R}^{N \times N}$  is denoted as  $V$ . The discrete gradient operator is a mapping  $\nabla : V \rightarrow Q$ , where  $Q = V \times V$ . For  $u \in V$ ,  $\nabla u$  is given by

$$(\nabla u)_{i,j} = ((\mathring{D}_x^+ u)_{i,j}, (\mathring{D}_y^+ u)_{i,j}),$$

with

$$\begin{aligned} (\mathring{D}_x^+ u)_{i,j} &= \begin{cases} u_{i,j+1} - u_{i,j}, & 1 \leq j \leq N-1 \\ u_{i,1} - u_{i,N}, & j = N \end{cases} \\ (\mathring{D}_y^+ u)_{i,j} &= \begin{cases} u_{i+1,j} - u_{i,j}, & 1 \leq i \leq N-1 \\ u_{1,j} - u_{N,j}, & i = N \end{cases} \end{aligned} \quad ,$$

where  $i, j = 1, \dots, N$ . Here we use  $\mathring{D}_x^+$  and  $\mathring{D}_y^+$  to denote forward difference operators with periodic boundary condition ( $u$  is periodically extended). Consequently FFT can be adopted in our algorithm.

We denote the usual inner product and Euclidean norm of  $V$  as  $(\cdot, \cdot)_V$  and  $\|\cdot\|_V$ , respectively. We also equip the space  $Q$  with inner product  $(\cdot, \cdot)_Q$  and norm  $\|\cdot\|_Q$ , which are defined as follows. For  $p = (p^1, p^2) \in Q$  and  $q = (q^1, q^2) \in Q$ ,

$$(p, q)_Q = (p^1, q^1)_V + (p^2, q^2)_V,$$

and

$$\|p\|_Q = \sqrt{(p, p)_Q}.$$

In addition, we mention that, at each pixel  $(i, j)$ ,

$$|p_{i,j}| = |(p_{i,j}^1, p_{i,j}^2)| = \sqrt{(p_{i,j}^1)^2 + (p_{i,j}^2)^2},$$

the usual Euclidean norm in  $\mathbb{R}^2$ . From the subscript  $i, j$ , one may regard  $|p_{i,j}|$  as pixel-by-pixel norm of  $p$ .

Using the inner products of  $V$  and  $Q$ , we can find the adjoint operator of  $-\nabla$ , i.e., the discrete divergence operator  $\text{div} : Q \rightarrow V$ . Given  $p = (p^1, p^2) \in Q$ , we have

$$(\text{div}p)_{i,j} = p_{i,j}^1 - p_{i,j-1}^1 + p_{i,j}^2 - p_{i-1,j}^2 = (\mathring{D}_x^- p^1)_{i,j} + (\mathring{D}_y^- p^2)_{i,j},$$

where  $\mathring{D}_x^-$  and  $\mathring{D}_y^-$  are backward difference operators with periodic boundary conditions  $p_{i,0}^1 = p_{i,N}^1$  and  $p_{0,j}^2 = p_{N,j}^2$ .

**3. The ROF model and some existing solvers.** Assume  $f \in V$  is an observed image and is degraded from the true image,  $u \in V$ , as follows

$$f = Ku + n, \quad (3.1)$$

where  $K : V \rightarrow V$  is a convolution operator, and  $n \in V$  is the random Gaussian noise (the most typical noise model). Image restoration aims at recovering  $u$  from  $f$ . Since the problem is usually ill-posed, we cannot directly solve  $u$  from (3.1). Regularization on the solution should be considered. One of the most basic and successful image regularization models is the ROF model [33], which reads

$$\min_{u \in V} \{F_{\text{rof}}(u) = R_{\text{rof}}(\nabla u) + \frac{\alpha}{2} \|Ku - f\|_V^2\}, \quad (3.2)$$

where

$$R_{\text{rof}}(\nabla u) = \text{TV}(u) = \sum_{1 \leq i,j \leq N} |(\nabla u)_{i,j}|, \quad (3.3)$$

is the total variation of  $u$ . Note here  $R_{\text{rof}}(\cdot)$  is regarded as a functional of  $\nabla u$ . In [33], the authors considered the image denoising problem ( $K = I$ ) and presented a gradient descent method to solve (3.2). Here the method is described for general  $K$ . The artificial time marching is introduced to the associated Euler-Lagrange equation (it is actually a system of ordinary differential equations since we are in discrete setting) as follows

$$\begin{aligned} u_t &= \text{div}\left(\frac{\nabla u}{\sqrt{|\nabla u|^2 + \beta}}\right) + \alpha K^*(f - Ku), \\ u(0) &= f \end{aligned}, \quad (3.4)$$

where  $\beta$  is a small positive number to avoid zero division and  $K^*$  is the  $L_2$  adjoint of  $K$ .

There are mainly two drawbacks for the gradient descent method (3.4). At first, it is an approximation of the original problem (3.2), since the regularity term  $R_{\text{rof}}(\nabla u)$  is smoothed and thus approximated to get (3.4). On the second, the method is slow due to strict constraints on the time step. The choice of  $\beta$  will effect both of these aspects. Larger the  $\beta$ , more efficient the scheme is, whereas worse the approximation will be. Therefore it is a tradeoff between the accuracy and the efficiency.

Many algorithms have been proposed to improve the gradient descent method, aiming to compute the solution of the ROF model (3.2) as efficiently and exactly as possible; see, e.g., dual methods [12, 11], split Bregman iteration [43, 22], as well as splitting and penalty based method [39, 40].

The difficulty to solve the ROF restoration model (3.2) is due to the non-differentiability of the total variation norm. By using an operator-splitting technique [21, 39, 40, 22], we can separate the calculation of the non-differentiable term and the squared 2-norm

term. Concretely, an auxiliary variable  $p \in Q$  is introduced for  $\nabla u$ . The model (3.2) is thus equivalent to

$$\begin{aligned} & \min_{u \in V, p \in Q} \{G_{\text{rof}}(u, p) = R_{\text{rof}}(p) + \frac{\alpha}{2} \|Ku - f\|_V^2\} \\ & \text{s.t.} \quad p = \nabla u \end{aligned}, \quad (3.5)$$

which is a constrained optimization problem.

In this paper, we make the following mild assumption

- The null spaces of  $\nabla$  and  $K$  have only 0 as common elements, i.e.,  $\text{Null}(\nabla) \cap \text{Null}(K) = \{0\}$ .

Under this assumption, the functional  $F_{\text{rof}}(u)$  in (3.2) is convex, proper, coercive, and continuous. According to the generalized Weierstrass theorem and Fermat's rule [19, 21], we have the following result.

**THEOREM 3.1.** *The problem (3.2) has at least one solution  $u$ , which satisfies*

$$0 \in \alpha K^*(Ku - f) - \text{div} \partial R_{\text{rof}}(\nabla u), \quad (3.6)$$

where  $\partial R_{\text{rof}}(\nabla u)$  is the sub-differential [19] of  $R_{\text{rof}}$  at  $\nabla u$ . Moreover, if  $\text{Null}(K) = \{0\}$ , the minimizer is unique.

In the following, we review some typical existing solvers for the ROF model.

**3.1. The CGM dual method.** In [12] Chan *et al* proposed a primal-dual method to solve the ROF model. They introduced a new variable  $\omega \in Q$  defined by

$$\omega_{i,j} = \frac{(\nabla u)_{i,j}}{|(\nabla u)_{i,j}|}, 1 \leq i, j \leq N \quad (3.7)$$

to the Euler-Lagrange equation of (3.2), to remove some of the singularity caused by the non-differentiability of the object functional. This yields the following primal-dual system:

$$\begin{aligned} -\text{div} \omega + \alpha K^*(Ku - f) &= 0 \\ \nabla u - \omega |\nabla u| &= 0 \end{aligned}, \quad (3.8)$$

where  $u$  and  $\omega$  are called primal and dual variables, respectively. The system is then approximated using a regularized TV norm (with some small positive  $\beta$ ) in numerical computation. Newton's linearization technique for both the primal and dual variables is adopted. As shown in [18], the primal-dual Newton's method is very efficient and the parameter  $\beta$  can be very close to 0.

**3.2. Chambolle's dual method.** Another work based on dual formulation with a different derivation is due to Chambolle [11]. In this method, the primal variable of the image data is expressed explicitly with the dual variable and only the dual variable is computed iteratively. However, the algorithm does not consider general operator  $K$  in (3.2). In the following we introduce Chambolle's method in our context (Note the difference on the boundary condition we used, and the slight difference between (3.2) and the model in [11] about the parameter  $\alpha$ ).

Denoting

$$S = \text{Closure}\{\text{div} \xi : \xi \in Q, |\xi_{i,j}| \leq 1, \forall 1 \leq i, j \leq N\}, \quad (3.9)$$

Chambolle [11] showed that the ROF restoration model (3.2) with  $K = I$  yields

$$u = f - \frac{1}{\alpha} \pi_S(\alpha f) = f - \pi_{\frac{\alpha}{\alpha}}(f), \quad (3.10)$$

where  $\pi_S(\cdot)$  is a nonlinear projection operator to  $S$ , which reads

$$\min_{\text{div}\xi} \{ \|\text{div}\xi - \cdot\|_V^2 : \xi \in Q, |\xi_{i,j}| \leq 1, \forall 1 \leq i, j \leq N \}. \quad (3.11)$$

From the Karush-Kuhn-Tucker (KKT) conditions and a careful observation, it was shown that  $\xi$  in the nonlinear projection satisfies

$$-(\nabla(\text{div}\xi - \alpha f))_{i,j} + \xi_{i,j} |(\nabla(\text{div}\xi - \alpha f))_{i,j}| = 0, \quad (3.12)$$

which allows a semi-implicit gradient descent algorithm to find  $\xi$ .

**3.3. Split Bregman iteration.** Recently, Bregman iteration and split Bregman iteration attract much attention in signal recovery and image processing community [7, 8, 22, 30, 43, 44, 47]. The basic idea is to transform a constrained optimization problem to a series of unconstrained problems. In each unconstrained problem, the object function is defined by the Bregman distance [3] of a convex function.

The Bregman distance of a convex functional  $J(u)$  is defined as the following (nonnegative) quantity

$$D_J^g(u, v) \equiv J(u) - J(v) - \langle g, u - v \rangle, \quad (3.13)$$

where  $g \in \partial J(v)$ , i.e., one of the sub-gradients of  $J$  at  $v$ .

When  $J(u)$  is a continuously differentiable functional, its sub-differential  $\partial J(v)$  has a single element for each  $v$ , and consequently the Bregman distance is unique. In this case the distance is just the difference at the point  $u$  between  $J(\cdot)$  and its first order approximation at the point  $v$ . For those non-differentiable functionals, the sub-differential may contain none or multiple values. Therefore, the Bregman distance between  $u$  and  $v$  can be ill-defined or multi-valued. However, this doesn't matter in Bregman distance based iterative algorithms since the algorithms automatically choose a unique sub-gradient in each iteration as long as the fidelity term for the constraints is differentiable (This condition holds usually). We also remind here that the Bregman distance of a functional is not a distance in the usual sense since, in general,  $D_J^g(u, v) \neq D_J^g(v, u)$  and the triangle inequality does not hold. See [30, 43] for more details.

To find the solution of the ROF model (3.2), or equivalently the constrained problem (3.5), split Bregman iteration solves a sequence of unconstrained problems with the form as

$$(u^k, p^k) = \arg \min_{u \in V, p \in Q} D_{G_{\text{rof}}}^{(g_u^{k-1}, g_p^{k-1})}((u, p), (u^{k-1}, p^{k-1})) + \frac{1}{2} \|p - \nabla u\|_Q^2, \quad (3.14)$$

where  $g_u^{k-1}$  and  $g_p^{k-1}$ , sometime written together to be  $(g_u^{k-1}, g_p^{k-1})$ , are the sub-gradients of  $G_{\text{rof}}$  at  $(u^{k-1}, p^{k-1})$  with respect to  $u$  and  $p$ , respectively. Taking the update of the sub-gradients into consideration, the iteration procedure is formulated as Algorithm 3.1. The computation of  $(u^k, p^k)$  in the algorithm is similar with Algorithm 4.2.

**4. Augmented Lagrangian method for the ROF model.** Augmented Lagrangian method [23, 31, 32] has many advantages over other methods such as penalty method [1], and has been successfully applied to nonlinear PDEs and mechanics [21]. In this section, we present the method for the ROF model, or equivalently the constrained problem (3.5). The details of the algorithms will be elaborated, followed by some convergence analysis.

---

**Algorithm 3.1** Split Bregman iteration for the ROF model
 

---

1. *Initialization:*  $u^{-1} = 0, p^{-1} = 0, g_u^{-1} = 0, g_p^{-1} = 0$ ;
2. For  $k=0, 1, 2, \dots$ : Compute  $(u^k, p^k)$  using (3.14), and update

$$\begin{aligned} g_u^k &= g_u^{k-1} - \operatorname{div}(p^k - \nabla u^k) \\ g_p^k &= g_p^{k-1} - (p^k - \nabla u^k) \end{aligned} \quad (3.15)$$


---

**4.1. Augmented Lagrangian method for the ROF model.** We first define the augmented Lagrangian functional for the constrained optimization problem (3.5) as follows:

$$\mathcal{L}_{\text{rof}}(v, q; \mu) = R_{\text{rof}}(q) + \frac{\alpha}{2} \|Kv - f\|_V^2 + (\mu, q - \nabla v)_Q + \frac{r}{2} \|q - \nabla v\|_Q^2, \quad (4.1)$$

where  $\mu \in Q$  is the Lagrange multiplier, and  $r$  is a positive constant. For the augmented Lagrangian method for (3.5), we consider the following saddle-point problem

$$\begin{aligned} \text{Find } & (u, p; \lambda) \in V \times Q \times Q, \\ \text{s.t. } & \mathcal{L}_{\text{rof}}(u, p; \mu) \leq \mathcal{L}_{\text{rof}}(u, p; \lambda) \leq \mathcal{L}_{\text{rof}}(v, q; \lambda), \forall (v, q; \mu) \in V \times Q \times Q. \end{aligned} \quad (4.2)$$

The relation between the saddle-point of problem (4.2) and the solution of (3.2) is stated in the following theorem.

**THEOREM 4.1.**  *$u \in V$  is a solution of (3.2) if and only if there exist  $p \in Q$  and  $\lambda \in Q$  such that  $(u, p; \lambda)$  is a solution of (4.2).*

**Proof** Suppose  $(u, p; \lambda)$  is a solution of (4.2). From the first inequality in (4.2), we have

$$p - \nabla u = 0. \quad (4.3)$$

The above relation, together with the second inequality in (4.2), shows

$$\begin{aligned} R_{\text{rof}}(\nabla u) + \frac{\alpha}{2} \|Ku - f\|_V^2 &\leq R_{\text{rof}}(q) + \frac{\alpha}{2} \|Kv - f\|_V^2 + (\lambda, q - \nabla v)_Q + \frac{r}{2} \|q - \nabla v\|_Q^2, \\ &\forall (v, q) \in V \times Q. \end{aligned} \quad (4.4)$$

Taking  $q = \nabla v$  in the above equation indicates that  $u$  is a solution of (3.2).

Conversely, we assume that  $u \in V$  is a solution of (3.2). We take  $p = \nabla u \in Q$ . From (3.6), there exists one  $\lambda \in \partial R_{\text{rof}}(\nabla u)$  such that  $\operatorname{div} \lambda = \alpha K^*(Ku - f)$ . We verify that  $(u, p; \lambda)$  is a saddle-point of  $\mathcal{L}_{\text{rof}}$ , i.e.,  $\mathcal{L}_{\text{rof}}(u, p; \mu) \leq \mathcal{L}_{\text{rof}}(u, p; \lambda) \leq \mathcal{L}_{\text{rof}}(v, q; \lambda), \forall (v, q; \mu) \in V \times Q \times Q$ . Since  $p = \nabla u$ , the first inequality holds. In the following we show  $\mathcal{L}_{\text{rof}}(u, p; \lambda) \leq \mathcal{L}_{\text{rof}}(v, q; \lambda), \forall (v, q) \in V \times Q$ . Since

$$\begin{aligned} \mathcal{L}_{\text{rof}}(v, q; \lambda) &= R_{\text{rof}}(q) + \frac{\alpha}{2} \|Kv - f\|_V^2 + (\lambda, q - \nabla v)_Q + \frac{r}{2} \|q - \nabla v\|_Q^2 \\ &= R_{\text{rof}}(q) + \frac{\alpha}{2} \|Kv - f\|_V^2 + \frac{r}{2} \|q - \nabla v + \frac{\lambda}{r}\|_Q^2 - \frac{1}{2r} \|\lambda\|_Q^2 \end{aligned}$$

is convex, proper, coercive, and continuous with respect to  $(v, q)$ ,  $\mathcal{L}_{\text{rof}}(v, q; \lambda)$  has a minimizer  $(\bar{v}, \bar{q})$  over  $V \times Q$ , which is characterized [19, 21] by

$$R_{\text{rof}}(q) - R_{\text{rof}}(\bar{q}) + (\lambda, q - \bar{q})_Q + r(\bar{q} - \nabla \bar{v}, q - \bar{q})_Q \geq 0, \forall q \in Q, \quad (4.5)$$

and

$$\frac{\alpha}{2}\|Kv-f\|_V^2 - \frac{\alpha}{2}\|K\bar{v}-f\|_V^2 + (\operatorname{div}\lambda, v-\bar{v})_V + r(\operatorname{div}(\bar{q}-\nabla\bar{v}), v-\bar{v})_V \geq 0, \forall v \in V. \quad (4.6)$$

It is straightforward to verify that  $(u, p)$  satisfies (4.5) and (4.6). This completes the proof.  $\blacksquare$

Theorem 4.1, together with Theorem 3.1, shows that the problem (4.2) has at least one solution and each  $u$  in the solutions solves the original problem (3.2). We then use an iterative algorithm to solve the saddle-point problem (4.2); see Algorithm 4.1.

---

**Algorithm 4.1** Augmented Lagrangian method for the ROF model

---

1. *Initialization:*  $\lambda^0 = 0$ ;
2. For  $k=0, 1, 2, \dots$ : compute  $(u^k, p^k)$  as an (approximate) minimizer of the augmented Lagrangian functional with the Lagrange multiplier  $\lambda^k$ , i.e.,

$$(u^k, p^k) \approx \arg \min_{(v, q) \in V \times Q} \mathcal{L}_{\text{rof}}(v, q; \lambda^k), \quad (4.7)$$

where  $\mathcal{L}_{\text{rof}}(v, q; \lambda^k)$  is defined in (4.1); and update

$$\lambda^{k+1} = \lambda^k + r(p^k - \nabla u^k). \quad (4.8)$$


---

We are now left the minimization problem (4.7) to address. One may notice the symbol  $\approx$  in this problem. This is because that, in general, it is difficult to find the minimizers  $u^k$  and  $p^k$  exactly in practical computation since  $v, q$  are coupled together. Usually, one separates the variable  $v$  and  $q$  and then uses an alternative minimization procedure [39, 40, 22] to solve (4.7), through which in practice one can only obtain the minimizer approximately. However, this does not affect the convergence of the whole algorithm 4.1. More details are as follows.

We separate (4.7) to be the following two sub-problems:

$$\min_{v \in V} \frac{\alpha}{2}\|Kv-f\|_V^2 - (\lambda^k, \nabla v)_Q + \frac{r}{2}\|q - \nabla v\|_Q^2, \quad (4.9)$$

for a given  $q$ , and

$$\min_{q \in Q} R_{\text{rof}}(q) + (\lambda^k, q)_Q + \frac{r}{2}\|q - \nabla v\|_Q^2, \quad (4.10)$$

for a given  $v$ .

Sub-problems (4.9) and (4.10) can be efficiently solved. For (4.9), the optimality condition gives a linear equation

$$\alpha K^*(Kv-f) + \operatorname{div}\lambda^k + r\operatorname{div}q - r\Delta v = 0,$$

by the periodic boundary condition we are using. It allows us to use Fourier transforms and thus an FFT implementation as done in [39, 40], which, to our best of knowledge, are the first papers using FFT in total variation minimization problems. Denoting  $\mathcal{F}(v)$  as the Fourier transform of  $v$ , we write the solution as follows

$$v = \mathcal{F}^{-1}\left(\frac{\alpha\mathcal{F}(K^*)\mathcal{F}(f) - \mathcal{F}(\overset{\circ}{D}_x^-)((\lambda^1)^k + rq^1) - \mathcal{F}(\overset{\circ}{D}_y^-)((\lambda^2)^k + rq^2)}{\alpha\mathcal{F}(K^*)\mathcal{F}(K) - r\mathcal{F}(\Delta)}\right), \quad (4.11)$$

where  $\lambda^k = ((\lambda^1)^k, (\lambda^2)^k)$  and  $q = (q^1, q^2)$ ; and Fourier transforms of operators such as  $K, \hat{D}_x^-, \hat{D}_y^-, \Delta = \hat{D}_x^- \hat{D}_x^+ + \hat{D}_y^- \hat{D}_y^+$  are regarded as the transforms of their corresponding convolution kernels. For (4.10), we actually have the following closed form solution [6][40]

$$q_{i,j} = \begin{cases} (1 - \frac{1}{r} \frac{1}{|w_{i,j}|}) w_{i,j}, & |w_{i,j}| > \frac{1}{r}, \\ 0, & |w_{i,j}| \leq \frac{1}{r}, \end{cases} \quad (4.12)$$

where

$$w = \nabla v - \frac{\lambda^k}{r}, \quad (4.13)$$

since we can reformulate it (by multiplying  $r$ ) to be

$$\min_{q \in Q} R_{\text{rof}}(rq) + \frac{1}{2} \|rq - (r\nabla v - \lambda^k)\|_Q^2.$$

We then iteratively and alternatively compute the  $v$  and  $q$  according to (4.11) and (4.12). It is with Gauss-Seidel flavor. The procedure is shown in Algorithm 4.2. Here

---

**Algorithm 4.2** Augmented Lagrangian method for the ROF model – solve the minimization problem (4.7)

---

- *Initialization:*  $u^{k,0} = u^{k-1}, p^{k,0} = p^{k-1}$ ;
  - *For*  $l = 0, 1, 2, \dots, L-1$ : Compute  $u^{k,l+1}$  from (4.11) for  $q = p^{k,l}$ ; and then compute  $p^{k,l+1}$  from (4.12) for  $v = u^{k,l+1}$ ;
  - $u^k = u^{k,L}, p^k = p^{k,L}$ .
- 

$L$  can be chosen using some convergence test techniques. In this paper, we simply set  $L = 1$ . In our experiments we found that with larger  $L$  ( $> 1$ ) the algorithm wastes the accuracy of the inner iteration and does not speed up dramatically the convergence of the whole algorithm. This has also been observed in [22], for the split Bregman method (which is equivalent to augmented Lagrangian method, as will be shown in the following).

**4.2. Convergence analysis.** We show some convergence results of the augmented Lagrangian method. We first give the convergence of Algorithm 4.2, and then present two convergence results for Algorithm 4.1 where the minimization problem (4.7) is computed by Algorithm 4.2 with full accuracy ( $L \rightarrow \infty$ ) and rough accuracy ( $L = 1$ ), respectively.

**THEOREM 4.2.** *The sequence  $\{(u^{k,l}, p^{k,l}) : l = 0, 1, 2, \dots\}$  generated by Algorithm 4.2 converges to a solution of the problem (4.7).*

**Proof** The proof is motivated by [40]. Here we just sketch the differences.

We define an operator  $S$  similarly with that in [40], such that (4.12) is reformulated as  $q = S(w)$ , where  $w$  is as in (4.13). Therefore the iterative scheme in Algorithm 4.2 is as follows

$$\begin{cases} u^{k,l+1} = (\nabla^* \nabla + \frac{\alpha}{r} K^* K)^{-1} (\nabla^* p^{k,l} + \nabla^* \frac{\lambda^k}{r} + \frac{\alpha}{r} K^* f), \\ p^{k,l+1} = S(\nabla u^{k,l+1} - \frac{\lambda^k}{r}), \end{cases} \quad (4.14)$$

where  $\nabla^* = -\text{div}$  is the adjoint operator of  $\nabla$ . Here we also mention the existence of  $(\nabla^* \nabla + \frac{\alpha}{r} K^* K)^{-1}$  for the assumption  $\text{Null}(\nabla) \cap \text{Null}(K) = \{0\}$ . We then define a



linear operator  $h : Q \rightarrow Q$  as

$$h(q) = \nabla(\nabla^* \nabla + \frac{\alpha}{r} K^* K)^{-1} (\nabla^* q + \nabla^* \frac{\lambda^k}{r} + \frac{\alpha}{r} K^* f) - \frac{\lambda^k}{r}. \quad (4.15)$$

It is straightforward to verify the non-expansiveness of  $h$  defined above.

Rewriting the iterative scheme (4.14) as

$$\begin{cases} u^{k,l+1} = (\nabla^* \nabla + \frac{\alpha}{r} K^* K)^{-1} (\nabla^* p^{k,l} + \nabla^* \frac{\lambda^k}{r} + \frac{\alpha}{r} K^* f), \\ p^{k,l+1} = S \circ h(p^{k,l}), \end{cases} \quad (4.16)$$

one can show the convergence via a similar argument in [40].  $\blacksquare$

In the following we give the convergence of Algorithm 4.1 where the minimization problem (4.7) is computed by Algorithm 4.2 with full accuracy ( $L \rightarrow \infty$ ) and rough accuracy ( $L = 1$ ), respectively. We should point out that the idea of our proofs follows the convergence proof in [21]. However, the convergence proof of  $(u^k, p^k)$  (see (4.18) and (4.36)) in [21] requires the uniform convexity of  $R_{\text{rof}}(p)$  (in our context) and thus cannot be directly applied to our case. In addition to modifying this part, we provide more details to make the proof clearer.

**THEOREM 4.3.** *Assume  $(u, p; \lambda)$  is a saddle-point of  $\mathcal{L}_{\text{rof}}(v, q; \mu)$ . Suppose that the minimization problem (4.7) is exactly solved in each iteration, i.e.,  $L \rightarrow \infty$  in Algorithm 4.2. Then the sequence  $(u^k, p^k; \lambda^k)$  generated by Algorithm 4.1 satisfies*

$$\begin{cases} \lim_{k \rightarrow \infty} G_{\text{rof}}(u^k, p^k) = G_{\text{rof}}(u, p), \\ \lim_{k \rightarrow \infty} \|p^k - \nabla u^k\|_Q = 0. \end{cases} \quad (4.17)$$

Since  $R_{\text{rof}}(p)$  is continuous, (4.17) indicates that  $u^k$  is a minimizing sequence of  $F_{\text{rof}}$ . If we further have  $\text{Null}(K) = \{0\}$ , then

$$\begin{cases} \lim_{k \rightarrow \infty} u^k = u, \\ \lim_{k \rightarrow \infty} p^k = p. \end{cases} \quad (4.18)$$

**Proof** Let us define  $\bar{u}^k, \bar{p}^k, \bar{\lambda}^k$  as

$$\bar{u}^k = u^k - u, \quad \bar{p}^k = p^k - p, \quad \bar{\lambda}^k = \lambda^k - \lambda.$$

Since  $(u, p; \lambda)$  is a saddle-point of  $\mathcal{L}_{\text{rof}}(v, q; \mu)$ , we have

$$\mathcal{L}_{\text{rof}}(u, p; \mu) \leq \mathcal{L}_{\text{rof}}(u, p; \lambda) \leq \mathcal{L}_{\text{rof}}(v, q; \lambda), \quad \forall (v, q; \mu) \in V \times Q \times Q. \quad (4.19)$$

From the first inequality of (4.19), we have  $p = \nabla u$ . This relationship, together with (4.8), indicates

$$\bar{\lambda}^{k+1} = \bar{\lambda}^k + r(\bar{p}^k - \nabla \bar{u}^k).$$

It then follows that

$$\|\bar{\lambda}^k\|_Q^2 - \|\bar{\lambda}^{k+1}\|_Q^2 = -2r(\bar{\lambda}^k, \bar{p}^k - \nabla \bar{u}^k)_Q - r^2 \|\bar{p}^k - \nabla \bar{u}^k\|_Q^2. \quad (4.20)$$

On the other hand, from the second inequality of (4.19),  $(u, p)$  is characterized by

$$\frac{\alpha}{2} \|Kv - f\|_V^2 - \frac{\alpha}{2} \|Ku - f\|_V^2 + (\text{div} \lambda, v - u)_V + r(\text{div}(p - \nabla u), v - u)_V \geq 0, \quad \forall v \in V, \quad (4.21)$$

$$R_{\text{rof}}(q) - R_{\text{rof}}(p) + (\lambda, q - p)_Q + r(p - \nabla u, q - p)_Q \geq 0, \forall q \in Q. \quad (4.22)$$

Similarly,  $(u^k, p^k)$  is characterized by

$$\frac{\alpha}{2} \|Kv - f\|_V^2 - \frac{\alpha}{2} \|Ku^k - f\|_V^2 + (\text{div} \lambda^k, v - u^k)_V + r(\text{div}(p^k - \nabla u^k), v - u^k)_V \geq 0, \forall v \in V, \quad (4.23)$$

$$R_{\text{rof}}(q) - R_{\text{rof}}(p^k) + (\lambda^k, q - p^k)_Q + r(p^k - \nabla u^k, q - p^k)_Q \geq 0, \forall q \in Q, \quad (4.24)$$

since  $(u^k, p^k)$  is the solution of (4.7). Taking  $v = u^k$  in (4.21) and  $v = u$  in (4.23), we obtain, by addition,

$$-(\bar{\lambda}^k, \nabla \bar{u}^k)_Q - r(\bar{p}^k - \nabla \bar{u}^k, \nabla \bar{u}^k)_Q \leq 0. \quad (4.25)$$

Similarly, we have

$$(\bar{\lambda}^k, \bar{p}^k)_Q + r(\bar{p}^k - \nabla \bar{u}^k, \bar{p}^k)_Q \leq 0, \quad (4.26)$$

by taking  $q = p^k$  in (4.22) and  $q = p$  in (4.24) and then addition. It then follows that

$$(\bar{\lambda}^k, \bar{p}^k - \nabla \bar{u}^k)_Q + r\|\bar{p}^k - \nabla \bar{u}^k\|_Q^2 \leq 0, \quad (4.27)$$

if we add (4.25) and (4.26) together.

By (4.27) and (4.20), we have

$$\|\bar{\lambda}^k\|_Q^2 - \|\bar{\lambda}^{k+1}\|_Q^2 \geq r^2 \|\bar{p}^k - \nabla \bar{u}^k\|_Q^2 \geq 0, \quad (4.28)$$

which implies

$$\begin{cases} \{\lambda^k : \forall k\} \text{ is bounded,} \\ \lim_{k \rightarrow \infty} \|p^k - \nabla u^k\|_Q = 0. \end{cases} \quad (4.29)$$

Moreover, the second inequality of (4.19) indicates

$$G_{\text{rof}}(u, p) \leq G_{\text{rof}}(u^k, p^k) + (\lambda, p^k - \nabla u^k)_Q + \frac{r}{2} \|p^k - \nabla u^k\|_Q^2. \quad (4.30)$$

If we take  $v = u$  in (4.23) and  $q = p$  in (4.24), we have, by addition,

$$G_{\text{rof}}(u, p) \geq G_{\text{rof}}(u^k, p^k) + (\lambda^k, p^k - \nabla u^k)_Q + r\|p^k - \nabla u^k\|_Q^2. \quad (4.31)$$

Using (4.29), we have

$$\liminf G_{\text{rof}}(u^k, p^k) \geq G_{\text{rof}}(u, p) \geq \limsup G_{\text{rof}}(u^k, p^k), \quad (4.32)$$

by taking  $\liminf$  in (4.30) and  $\limsup$  in (4.31). Hence we complete the proof of (4.17).

In the following we show (4.18) if  $\text{Null}(K) = \{0\}$  holds. Since  $(u, p; \lambda)$  is a saddle-point of  $\mathcal{L}_{\text{rof}}(v, q; \mu)$ , we have

$$-\lambda \in \partial R_{\text{rof}}(p), \quad (4.33)$$

$$\text{div} \lambda = -\alpha K^*(Ku - f), \quad (4.34)$$

where  $\partial R_{\text{rof}}(p)$  is the sub-differential of  $R_{\text{rof}}$  at  $p$ . Then, we deduce

$$\begin{aligned}
 & G_{\text{rof}}(u^k, p^k) + (\lambda, p^k - \nabla u^k)_Q \\
 \geq & R_{\text{rof}}(p) - (\lambda, p^k - p)_Q + \frac{\alpha}{2} \|Ku^k - f\|_V^2 + (\lambda, p^k - \nabla u^k)_Q \\
 = & R_{\text{rof}}(p) + \frac{\alpha}{2} \|Ku^k - f\|_V^2 + (\lambda, \nabla u - \nabla u^k)_Q \\
 \geq & R_{\text{rof}}(p) + \frac{\alpha}{2} \|K\frac{u^k + u}{2} - f\|_V^2 + \alpha(K^*(K\frac{u^k + u}{2} - f), \frac{u^k - u}{2})_V + (\lambda, \nabla u - \nabla u^k)_Q \\
 = & R_{\text{rof}}(p) + \frac{\alpha}{2} \|Ku - f\|_V^2 + \frac{\alpha}{2} \|K\frac{u^k + u}{2} - f\|_V^2 - \frac{\alpha}{2} \|Ku - f\|_V^2 \\
 & + \alpha(K^*(K\frac{u^k + u}{2} - f), \frac{u^k - u}{2})_V + (\lambda, \nabla u - \nabla u^k)_Q \\
 = & R_{\text{rof}}(p) + \frac{\alpha}{2} \|Ku - f\|_V^2 + \frac{\alpha}{2} \|K\frac{u^k + u}{2} - f\|_V^2 - \frac{\alpha}{2} \|Ku - f\|_V^2 \\
 & + \alpha(K\frac{u^k + u}{2} - f, K\frac{u^k - u}{2})_V + \alpha(Ku - f, K(u - u^k))_V \\
 = & G_{\text{rof}}(u, p) + \frac{3}{8}\alpha \|K(u^k - u)\|_V^2,
 \end{aligned}$$

from which we obtain

$$\lim_{k \rightarrow \infty} \|K(u^k - u)\|_V = 0,$$

according to (4.17). If  $\text{Null}(K) = \{0\}$  holds, it follows that

$$\lim_{k \rightarrow \infty} u^k = u.$$

This result, together with the second equation in (4.17), yields

$$\lim_{k \rightarrow \infty} p^k = \nabla u = p,$$

which completes the proof.  $\blacksquare$

**THEOREM 4.4.** *Assume  $(u, p; \lambda)$  is a saddle-point of  $\mathcal{L}_{\text{rof}}(v, q; \mu)$ . Suppose that the minimization problem (4.7) is roughly solved in each iteration, i.e., with  $L = 1$  in Algorithm 4.2. Then the sequence  $(u^k, p^k; \lambda^k)$  generated by Algorithm 4.1 satisfies*

$$\begin{cases} \lim_{k \rightarrow \infty} G_{\text{rof}}(u^k, p^k) = G_{\text{rof}}(u, p), \\ \lim_{k \rightarrow \infty} \|p^k - \nabla u^k\|_Q = 0. \end{cases} \quad (4.35)$$

Since  $R_{\text{rof}}(p)$  is continuous, (4.35) indicates that  $u^k$  is a minimizing sequence of  $F_{\text{rof}}$ . If we further have  $\text{Null}(K) = \{0\}$ , then

$$\begin{cases} \lim_{k \rightarrow \infty} u^k = u, \\ \lim_{k \rightarrow \infty} p^k = p. \end{cases} \quad (4.36)$$

**Proof** Again we define the following errors

$$\bar{u}^k = u^k - u, \quad \bar{p}^k = p^k - p, \quad \bar{\lambda}^k = \lambda^k - \lambda.$$

In this case, (4.20) still holds, which is presented as follows

$$\|\bar{\lambda}^k\|_Q^2 - \|\bar{\lambda}^{k+1}\|_Q^2 = -2r(\bar{\lambda}^k, \bar{p}^k - \nabla \bar{u}^k)_Q - r^2\|\bar{p}^k - \nabla \bar{u}^k\|_Q^2. \quad (4.37)$$

Since  $(u, p; \lambda)$  is a saddle-point of  $\mathcal{L}_{\text{rof}}(v, q; \mu)$ ,  $(u, p)$  is characterized by

$$\frac{\alpha}{2}\|Kv - f\|_V^2 - \frac{\alpha}{2}\|Ku - f\|_V^2 + (\text{div} \lambda, v - u)_V + r(\text{div}(p - \nabla u), v - u)_V \geq 0, \forall v \in V, \quad (4.38)$$

$$R_{\text{rof}}(q) - R_{\text{rof}}(p) + (\lambda, q - p)_Q + r(p - \nabla u, q - p)_Q \geq 0, \forall q \in Q. \quad (4.39)$$

Similarly, by the construction of  $(u^k, p^k)$  (Algorithm 4.2 with  $L = 1$ ), we have

$$\frac{\alpha}{2}\|Kv - f\|_V^2 - \frac{\alpha}{2}\|Ku^k - f\|_V^2 + (\text{div} \lambda^k, v - u^k)_V + r(\text{div}(p^{k-1} - \nabla u^k), v - u^k)_V \geq 0, \forall v \in V, \quad (4.40)$$

$$R_{\text{rof}}(q) - R_{\text{rof}}(p^k) + (\lambda^k, q - p^k)_Q + r(p^k - \nabla u^k, q - p^k)_Q \geq 0, \forall q \in Q. \quad (4.41)$$

Taking  $v = u^k$  in (4.38),  $v = u$  in (4.40) and  $q = p^k$  in (4.39), as well as  $q = p$  in (4.41), we obtain, after addition,

$$(\bar{\lambda}^k, \bar{p}^k - \nabla \bar{u}^k)_Q + r\|\bar{p}^k - \nabla \bar{u}^k\|_Q^2 + r(\nabla \bar{u}^k, \bar{p}^k - \bar{p}^{k-1})_Q \leq 0. \quad (4.42)$$

It then follows from (4.37) and (4.42) that

$$\|\bar{\lambda}^k\|_Q^2 - \|\bar{\lambda}^{k+1}\|_Q^2 \geq r^2\|\bar{p}^k - \nabla \bar{u}^k\|_Q^2 + 2r^2(\nabla \bar{u}^k, \bar{p}^k - \bar{p}^{k-1})_Q. \quad (4.43)$$

In the following we estimate  $(\nabla \bar{u}^k, \bar{p}^k - \bar{p}^{k-1})_Q$  in (4.43). We have

$$\begin{aligned} (\nabla \bar{u}^k, \bar{p}^k - \bar{p}^{k-1})_Q &= (\nabla \bar{u}^k - \nabla \bar{u}^{k-1}, \bar{p}^k - \bar{p}^{k-1})_Q \\ &\quad + (\nabla \bar{u}^{k-1}, \bar{p}^k - \bar{p}^{k-1})_Q + (\bar{p}^{k-1}, \bar{p}^k - \bar{p}^{k-1})_Q. \end{aligned} \quad (4.44)$$

On the other hand, by the construction of  $p^{k-1}$  (from  $u^{k-1}$ ), it follows that

$$R_{\text{rof}}(q) - R_{\text{rof}}(p^{k-1}) + (\lambda^{k-1}, q - p^{k-1})_Q + r(p^{k-1} - \nabla u^{k-1}, q - p^{k-1})_Q \geq 0, \forall q \in Q. \quad (4.45)$$

Taking  $q = p^{k-1}$  in (4.41) and  $q = p^k$  in (4.45), we obtain, by addition,

$$r\|\bar{p}^k - \bar{p}^{k-1}\|_Q^2 + (\bar{p}^k - \bar{p}^{k-1}, \bar{\lambda}^k - \bar{\lambda}^{k-1})_Q - r(\bar{p}^k - \bar{p}^{k-1}, \nabla \bar{u}^k - \nabla \bar{u}^{k-1})_Q \leq 0. \quad (4.46)$$

Since

$$\bar{\lambda}^k - \bar{\lambda}^{k-1} = \lambda^k - \lambda^{k-1} = r(\bar{p}^{k-1} - \nabla \bar{u}^{k-1}),$$

we have

$$(\bar{p}^k - \bar{p}^{k-1}, \nabla \bar{u}^k - \nabla \bar{u}^{k-1})_Q + (\bar{p}^k - \bar{p}^{k-1}, \nabla \bar{u}^{k-1} - \bar{p}^{k-1})_Q \geq \|\bar{p}^k - \bar{p}^{k-1}\|_Q^2. \quad (4.47)$$

according to (4.46). (4.44) and (4.47), together with the following identity

$$(\bar{p}^{k-1}, \bar{p}^k - \bar{p}^{k-1})_Q = \frac{1}{2}(\|\bar{p}^k\|_Q^2 - \|\bar{p}^{k-1}\|_Q^2 - \|\bar{p}^k - \bar{p}^{k-1}\|_Q^2),$$

imply

$$(\nabla \bar{u}^k, \bar{p}^k - \bar{p}^{k-1})_Q \geq \frac{1}{2} (\|\bar{p}^k\|_Q^2 - \|\bar{p}^{k-1}\|_Q^2 + \|\bar{p}^k - \bar{p}^{k-1}\|_Q^2). \quad (4.48)$$

We then obtain from (4.43) and (4.48) that

$$\|\bar{\lambda}^k\|_Q^2 + r^2 \|\bar{p}^{k-1}\|_Q^2 - (\|\bar{\lambda}^{k+1}\|_Q^2 + r^2 \|\bar{p}^k\|_Q^2) \geq r^2 \|\bar{p}^k - \nabla \bar{u}^k\|_Q^2 + r^2 \|\bar{p}^k - \bar{p}^{k-1}\|_Q^2. \quad (4.49)$$

(4.49) indicates

$$\left\{ \begin{array}{l} \{\lambda^k : \forall k\}, \{p^k : \forall k\}, \text{ and } \{\nabla u^k : \forall k\} \text{ are bounded,} \\ \lim_{k \rightarrow \infty} \|p^k - \nabla u^k\|_Q = 0, \\ \lim_{k \rightarrow \infty} \|p^k - p^{k-1}\|_Q = 0. \end{array} \right. \quad (4.50)$$

On the other hand, since  $(u, p; \lambda)$  is a saddle-point of  $\mathcal{L}_{\text{rof}}(v, q; \mu)$ , we have

$$G_{\text{rof}}(u, p) \leq G_{\text{rof}}(u^k, p^k) + (\lambda, p^k - \nabla u^k)_Q + \frac{r}{2} \|p^k - \nabla u^k\|_Q^2. \quad (4.51)$$

If we take  $v = u$  in (4.40) and  $q = p$  in (4.41), we have, by addition,

$$G_{\text{rof}}(u, p) \geq G_{\text{rof}}(u^k, p^k) + (\lambda^k, p^k - \nabla u^k)_Q + r \|p^k - \nabla u^k\|_Q^2 + r (\nabla \bar{u}^k, \bar{p}^k - \bar{p}^{k-1})_Q. \quad (4.52)$$

Using (4.50), we have

$$\liminf G_{\text{rof}}(u^k, p^k) \geq G_{\text{rof}}(u, p) \geq \limsup G_{\text{rof}}(u^k, p^k), \quad (4.53)$$

by taking  $\liminf$  in (4.51) and  $\limsup$  in (4.52). This completes the proof of (4.35). Starting from (4.35), one can verify (4.36) in a similar way as in the proof of Theorem 4.3.  $\blacksquare$

Similar results as Theorem 4.4 were obtained in [36] by reformulating the algorithm to be Douglas-Rachford splitting on the dual problem, and also in [8].

We remind that the operator  $K$  is invertible in many cases, e.g., image denoising where  $K = I$ , and most of image deblurring problems (although the condition number of the blur kernel may be very bad). In these cases, Theorem 4.3 and 4.4 imply the convergence of the sequence  $\{u^k\}$  (either  $L \rightarrow \infty$  or  $L = 1$  in Algorithm 4.2) to the unique solution of the problem.

**5. Relations between augmented Lagrangian method and dual methods as well as split Bregman iteration for the ROF model.** In this section, we show that dual methods such as CGM [12] and Chambolle's [11] for the ROF model are closely connected to the augmented Lagrangian method. Also, we explain that split Bregman iteration [22] is equivalent to Algorithm 4.1.

For the saddle-point problem (4.2), we have the following optimality condition

$$\partial_v \mathcal{L}_{\text{rof}}(v, q; \mu)|_{(u, p; \lambda)} = \alpha K^*(Ku - f) + \text{div} \lambda + r \text{div}(p - \nabla u) = 0, \quad (5.1)$$

$$\partial_q \mathcal{L}_{\text{rof}}(v, q; \mu)|_{(u, p; \lambda)} = \partial R_{\text{rof}}(p) + \lambda + r(p - \nabla u) \ni 0, \quad (5.2)$$

$$\partial_\mu \mathcal{L}_{\text{rof}}(v, q; \mu)|_{(u, p; \lambda)} = p - \nabla u = 0, \quad (5.3)$$

where  $\partial R_{\text{rof}}(p)$  is the sub-differential of  $R_{\text{rof}}$  at  $p$ .

It is definitely true that various techniques such as Newton's and quasi Newton's linearizations can be applied to the above system of optimality condition to solve the saddle-point problem. Actually the optimality condition can be simplified to the CGM and Chambolle's dual methods as discussed in the following.

**5.1. Connection to the CGM dual method.** We show how to obtain the CGM dual method from augmented Lagrangian method. Using (5.3), we get  $p = \nabla u$ , which gives

$$\lambda_{i,j} = \begin{cases} -\frac{(\nabla u)_{i,j}}{|(\nabla u)_{i,j}|}, & \text{if } |(\nabla u)_{i,j}| \neq 0, \\ g \in \mathbb{R}^2, |g| \leq 1, & \text{if } |(\nabla u)_{i,j}| = 0. \end{cases} \quad (5.4)$$

from (5.2). Therefore, the multiplier  $\lambda$  is just the dual variable of the CGM method with a different sign. We then can reformulate the system of (5.1), (5.2), and (5.3) to be

$$\begin{aligned} \operatorname{div} \lambda + \alpha K^*(Ku - f) &= 0 \\ \nabla u + \lambda |\nabla u| &= 0 \end{aligned} \quad (5.5)$$

which is just the primal-dual system in [12] if  $-\lambda$  is replaced with  $\omega$ .

**5.2. Connection to Chambolle's dual method.** In the following we show how Chambolle's algorithm is connected to the augmented Lagrangian method. Compared to the derivation in [11], this is another way to obtain the dual method. From the system of (5.1), (5.2), and (5.3), we first eliminate the  $p$  variable to obtain (5.5), and then ulteriorly eliminate the  $\lambda$  variable to get  $u$  as following (assume  $\operatorname{Null}(K) = \{0\}$ )

$$u = (\alpha K^* K)^{-1}(\alpha K^* f - \operatorname{div} \lambda). \quad (5.6)$$

This yields the equation for the dual variable as:

$$\nabla((K^* K)^{-1}(\alpha K^* f - \operatorname{div} \lambda)) + \lambda |\nabla((K^* K)^{-1}(\alpha K^* f - \operatorname{div} \lambda))| = 0. \quad (5.7)$$

For image denoising problems where  $K = I$ , (5.7) and (5.6) are just the equations used by Chambolle in [11] to solve the dual variable and recover the primal variable  $u$ , respectively. (5.7) for the dual variable in [11] was obtained through a key result from the optimization theory. We deduce the same equation naturally from augmented Lagrangian method. In addition, (5.6) and (5.7) are formulations for general  $K$ . We should mention that  $K$  is sometimes compact and thus the condition number of  $K^* K$  is very bad. In this case the algorithm is not as efficient as expected.

**5.3. Connection to split Bregman iteration.** The split Bregman iteration is equivalent to augmented Lagrangian method [38, 36, 20]. Considering the zero initialization for the sub-gradients and the Lagrange multiplier and letting

$$(g_u^{k-1}, g_p^{k-1}) = -(\operatorname{div} \lambda^k, \lambda^k) \quad (5.8)$$

for  $k = 0, 1, 2, \dots$ , we have

$$\begin{aligned} (u^k, p^k) &= \arg \min_{u,p} D_{G_{\operatorname{rof}}}^{(g_u^{k-1}, g_p^{k-1})}((u, p), (u^{k-1}, p^{k-1})) + \frac{1}{2} \|p - \nabla u\|_Q^2 \\ &= \arg \min_{u,p} R_{\operatorname{rof}}(p) + \frac{\alpha}{2} \|Ku - f\|_V^2 + (u, \operatorname{div} \lambda^k)_V + (\lambda^k, p)_Q + \frac{1}{2} \|p - \nabla u\|_Q^2 \\ &= \arg \min_{u,p} R_{\operatorname{rof}}(p) + \frac{\alpha}{2} \|Ku - f\|_V^2 - (\lambda^k, \nabla u)_Q + (\lambda^k, p)_Q + \frac{1}{2} \|p - \nabla u\|_Q^2 \\ &= \arg \min_{u,p} \mathcal{L}_{\operatorname{rof}}(u, p; \lambda^k), \end{aligned}$$

indicating the equivalence of the solutions of Bregman iteration and augmented Lagrangian method with  $r = 1$ , if the sub-problems in these two methods are solved identically. In the context of compressive sensing, this equivalence has been pointed out in [43].

**6. Extension to vectorial TV model.** In this section, we extend our method and observations to vectorial TV restoration model. Let us denote in general an  $M$ -channel image by  $\mathbf{u} = (u_1, u_2, \dots, u_M)$ , where  $u_m \in V, \forall m = 1, 2, \dots, M$ . The intensity at pixel  $(i, j)$  is thus multi-valued, say,  $\mathbf{u}_{i,j} = ((u_1)_{i,j}, (u_2)_{i,j}, \dots, (u_M)_{i,j})$ . If  $M = 3$ , one gets usual color models such as RGB.

For convenience of description, we introduce the following notation

$$\begin{aligned}\mathbf{V} &= \underbrace{V \times V \times \dots \times V}_M, \\ \mathbf{Q} &= \underbrace{Q \times Q \times \dots \times Q}_M.\end{aligned}$$

Hence an  $M$ -channel image  $\mathbf{u}$  is an element of  $\mathbf{V}$ , and its gradient  $\nabla \mathbf{u} = (\nabla u_1, \nabla u_2, \dots, \nabla u_M)$  is an element of  $\mathbf{Q}$ . The usual inner products and norms in  $\mathbf{V}$  and  $\mathbf{Q}$  are as follows:

$$\begin{aligned}(\mathbf{u}, \mathbf{v})_{\mathbf{V}} &= \sum_{1 \leq m \leq M} (u_m, v_m)_V, \quad \|\mathbf{u}\|_{\mathbf{V}} = \sqrt{(\mathbf{u}, \mathbf{u})_{\mathbf{V}}}; \\ (\mathbf{p}, \mathbf{q})_{\mathbf{Q}} &= \sum_{1 \leq m \leq M} (p_m, q_m)_Q, \quad \|\mathbf{p}\|_{\mathbf{Q}} = \sqrt{(\mathbf{p}, \mathbf{p})_{\mathbf{Q}}}.\end{aligned}$$

For  $\mathbf{u} \in \mathbf{V}$  and  $\mathbf{p} \in \mathbf{Q}$ , we also define the following pixel-by-pixel norms

$$|\mathbf{u}_{i,j}| = \sqrt{\sum_{1 \leq m \leq M} (u_m)_{i,j}^2}$$

and

$$|\mathbf{p}_{i,j}| = \sqrt{\sum_{1 \leq m \leq M} |(p_m)_{i,j}|^2}$$

at each pixel  $(i, j)$ .

We consider the following vector-valued image restoration problem:

$$\min_{\mathbf{u} \in \mathbf{V}} \{F_{\text{vTV}}(\mathbf{u}) = R_{\text{vTV}}(\nabla \mathbf{u}) + \frac{\alpha}{2} \|\mathbf{K}\mathbf{u} - \mathbf{f}\|_{\mathbf{V}}^2\}, \quad (6.1)$$

where

$$R_{\text{vTV}}(\nabla \mathbf{u}) = \text{TV}(\mathbf{u}) = \sum_{1 \leq i, j \leq N} \sqrt{\sum_{1 \leq m \leq M} |(\nabla u_m)_{i,j}|^2} \quad (6.2)$$

is the vectorial TV norm [34, 4] (see [2] for some other choices), and  $\mathbf{f} = (f_1, f_2, \dots, f_M) \in \mathbf{V}$  is an observed image,  $\mathbf{K} = (K_{i,j})_{M \times M} : \mathbf{V} \rightarrow \mathbf{V}$  is the blur operator. The diagonal elements of  $\mathbf{K}$  denote within channel blur whereas the off-diagonal elements describe cross channel blur. Similarly with the ROF model, here we make the following assumption

- $\text{Null}(\nabla) \cap \text{Null}(\mathbf{K}) = \{0\}$ .

Under this assumption, the functional  $F_{\text{vTV}}(\mathbf{u})$  in (6.1) is convex, proper, coercive, and continuous. Therefore we have:

**THEOREM 6.1.** *The problem (6.1) has at least one solution  $\mathbf{u}$ , which satisfies*

$$0 \in \alpha \mathbf{K}^*(\mathbf{K}\mathbf{u} - \mathbf{f}) - \text{div} \partial R_{\text{vTV}}(\nabla \mathbf{u}), \quad (6.3)$$

where  $\partial R_{\text{vTV}}(\nabla \mathbf{u})$  is the sub-differential of  $R_{\text{vTV}}$  at  $\nabla \mathbf{u}$ . Moreover, if  $\text{Null}(\mathbf{K}) = \{0\}$ , the minimizer is unique.

By introducing a new variable  $\mathbf{p} = (p_1, p_2, \dots, p_M) \in \mathbf{Q}$ , the minimization problem (6.1) is equivalent to the following constrained optimization problem

$$\begin{aligned} \min_{\mathbf{u} \in \mathbf{V}, \mathbf{p} \in \mathbf{Q}} \{ & G_{\text{vTV}}(\mathbf{u}, \mathbf{p}) = R_{\text{vTV}}(\mathbf{p}) + \frac{\alpha}{2} \|\mathbf{K}\mathbf{u} - \mathbf{f}\|_{\mathbf{V}}^2 \}, \\ \text{s.t.} \quad & \mathbf{p} = \nabla \mathbf{u}. \end{aligned} \quad (6.4)$$

**6.1. Augmented Lagrangian method.** Here we present augmented Lagrangian method for the restoration problem (6.1), or equivalently (6.4). We first define augmented Lagrangian functional as

$$\mathcal{L}_{\text{vTV}}(\mathbf{v}, \mathbf{q}; \mu) = R_{\text{vTV}}(\mathbf{q}) + \frac{\alpha}{2} \|\mathbf{K}\mathbf{v} - \mathbf{f}\|_{\mathbf{V}}^2 + (\mu, \mathbf{q} - \nabla \mathbf{v})_{\mathbf{Q}} + \frac{r}{2} \|\mathbf{q} - \nabla \mathbf{v}\|_{\mathbf{Q}}^2, \quad (6.5)$$

where  $\mu \in \mathbf{Q}$  is the multiplier. Augmented Lagrangian method aims at solving the following saddle-point problem:

$$\begin{aligned} \text{Find } & (\mathbf{u}, \mathbf{p}; \lambda) \in \mathbf{V} \times \mathbf{Q} \times \mathbf{Q}, \\ \text{s.t. } & \mathcal{L}_{\text{vTV}}(\mathbf{u}, \mathbf{p}; \mu) \leq \mathcal{L}_{\text{vTV}}(\mathbf{u}, \mathbf{p}; \lambda) \leq \mathcal{L}_{\text{vTV}}(\mathbf{v}, \mathbf{q}; \lambda), \forall (\mathbf{v}, \mathbf{q}; \mu) \in \mathbf{V} \times \mathbf{Q} \times \mathbf{Q}. \end{aligned} \quad (6.6)$$

Similarly to Theorem 4.1, we have the following result.

**THEOREM 6.2.**  $\mathbf{u} \in \mathbf{V}$  is a solution of (6.1) if and only if there exist  $\mathbf{p} \in \mathbf{Q}$  and  $\lambda \in \mathbf{Q}$  such that  $(\mathbf{u}, \mathbf{p}; \lambda)$  is a solution of (6.6).

We use an iterative procedure as described in Algorithm 6.1 to solve the problem (6.6). Again, one may see the  $\approx$  in (6.7). This is because that the minimization problem (6.7) has two coupled variables and hence difficult to be solved exactly (see the following argument).

---

**Algorithm 6.1** Augmented Lagrangian method for vectorial TV model

---

1. *Initialization:*  $\lambda^0 = 0$ ;
2. For  $k = 0, 1, 2, \dots$ : Compute  $(\mathbf{u}^k, \mathbf{p}^k)$  from

$$(\mathbf{u}^k, \mathbf{p}^k) \approx \arg \min_{(\mathbf{v}, \mathbf{q}) \in (\mathbf{V}, \mathbf{Q})} \mathcal{L}_{\text{vTV}}(\mathbf{v}, \mathbf{q}; \lambda^k), \quad (6.7)$$

and update

$$\lambda^{k+1} = \lambda^k + r(\mathbf{p}^k - \nabla \mathbf{u}^k). \quad (6.8)$$


---

As for the minimization problem (6.7), we separate it into the following two sub-problems:

$$\min_{\mathbf{v}} \frac{\alpha}{2} \|\mathbf{K}\mathbf{v} - \mathbf{f}\|_{\mathbf{V}}^2 - (\lambda^k, \nabla \mathbf{v})_{\mathbf{Q}} + \frac{r}{2} \|\mathbf{q} - \nabla \mathbf{v}\|_{\mathbf{Q}}^2, \quad (6.9)$$

for a given  $\mathbf{q}$ , and

$$\min_{\mathbf{q}} R_{\text{vTV}}(\mathbf{q}) + (\lambda^k, \mathbf{q})_{\mathbf{Q}} + \frac{r}{2} \|\mathbf{q} - \nabla \mathbf{v}\|_{\mathbf{Q}}^2, \quad (6.10)$$

for a given  $\mathbf{v}$ .



Applying Fourier transforms to the optimality condition of the sub-problem (6.9), we have

$$\alpha\mathcal{F}(\mathbf{K}^*)\mathcal{F}(\mathbf{K})\mathcal{F}(\mathbf{v}) - r\mathcal{F}(\Delta)\mathcal{F}(\mathbf{v}) = \alpha\mathcal{F}(\mathbf{K}^*)\mathcal{F}(\mathbf{f}) - \mathcal{F}(\text{div})\mathcal{F}(\lambda^k) - r\mathcal{F}(\text{div})\mathcal{F}(\mathbf{q}), \quad (6.11)$$

from which  $\mathcal{F}(\mathbf{v})$  can be found and then  $\mathbf{v}$  via an inverse Fourier transform. Here applying Fourier transform to a matrix or a vector is regarded as applying Fourier transforms to its components, e.g.,  $\mathcal{F}(\mathbf{v}) = (\mathcal{F}(v_1), \mathcal{F}(v_2), \dots, \mathcal{F}(v_M))$ ,  $\mathcal{F}(\text{div})\mathcal{F}(\lambda^k) = (\mathcal{F}(\text{div})\mathcal{F}(\lambda_1^k), \mathcal{F}(\text{div})\mathcal{F}(\lambda_2^k), \dots, \mathcal{F}(\text{div})\mathcal{F}(\lambda_M^k))$ ,  $\mathcal{F}(\mathbf{K}) = (\mathcal{F}(K_{i,j}))_{M \times M}$ . Thus one needs to solve a system of linear algebraic equations of  $\mathcal{F}(\mathbf{v}) = (\mathcal{F}(v_1), \mathcal{F}(v_2), \dots, \mathcal{F}(v_M))$  since cross blurs exist in general. In a special case without cross blurs, say, the blur kernel matrix  $\mathbf{K}$  is a diagonal matrix, the  $\mathcal{F}(\mathbf{v})$  can be calculated component by component. The sub-problem (6.10) has the following closed form solution

$$\mathbf{q}_{i,j} = \begin{cases} (1 - \frac{1}{r} \frac{1}{|\mathbf{w}_{i,j}|})\mathbf{w}_{i,j}, & |\mathbf{w}_{i,j}| > \frac{1}{r}, \\ 0, & |\mathbf{w}_{i,j}| \leq \frac{1}{r}, \end{cases} \quad (6.12)$$

where

$$\mathbf{w} = \nabla \mathbf{v} - \frac{\lambda^k}{r}. \quad (6.13)$$

We then have an iterative procedure to alternatively compute the  $\mathbf{v}$  and  $\mathbf{q}$  according to (6.11) (6.12); see Algorithm 6.2. Here  $L$  can be chosen using some convergence

---

**Algorithm 6.2** Augmented Lagrangian method for vectorial TV model – solve the minimization problem (6.7)

---

- *Initialization:*  $\mathbf{u}^{k,0} = \mathbf{u}^{k-1}$ ,  $\mathbf{p}^{k,0} = \mathbf{p}^{k-1}$ ;
  - *For*  $l = 0, 1, 2, \dots, L - 1$ : Compute  $\mathbf{u}^{k,l+1}$  from (6.11) for  $\mathbf{q} = \mathbf{p}^{k,l}$ ; and then compute  $\mathbf{p}^{k,l+1}$  from (6.12) for  $\mathbf{v} = \mathbf{u}^{k,l+1}$ ;
  - $\mathbf{u}^k = \mathbf{u}^{k,L}$ ,  $\mathbf{p}^k = \mathbf{p}^{k,L}$ .
- 

test techniques and usually simply set to be 1.

In the following we present some convergence results without giving proofs. They are straightforward generalizations of Theorem 4.2, 4.3 and 4.4.

**THEOREM 6.3.** *The sequence  $\{(\mathbf{u}^{k,l}, \mathbf{p}^{k,l}) : l = 0, 1, 2, \dots\}$  generated by Algorithm 6.2 converges to a solution of the problem (6.7).*

**THEOREM 6.4.** *Assume  $(\mathbf{u}, \mathbf{p}; \lambda)$  is a saddle-point of  $\mathcal{L}_{\text{vtv}}(\mathbf{v}, \mathbf{q}; \mu)$ . Suppose that the minimization problem (6.7) is exactly solved in each iteration, i.e.,  $L \rightarrow \infty$  in Algorithm 6.2. Then the sequence  $(\mathbf{u}^k, \mathbf{p}^k; \lambda^k)$  generated by Algorithm 6.1 satisfies*

$$\begin{cases} \lim_{k \rightarrow \infty} G_{\text{vtv}}(\mathbf{u}^k, \mathbf{p}^k) = G_{\text{vtv}}(\mathbf{u}, \mathbf{p}), \\ \lim_{k \rightarrow \infty} \|\mathbf{p}^k - \nabla \mathbf{u}^k\|_{\mathbf{Q}} = 0. \end{cases} \quad (6.14)$$

Since  $R_{\text{vtv}}(\mathbf{p})$  is continuous, (6.14) indicates that  $\mathbf{u}^k$  is a minimizing sequence of  $F_{\text{vtv}}$ . If we further have  $\text{Null}(\mathbf{K}) = \{0\}$ , then

$$\begin{cases} \lim_{k \rightarrow \infty} \mathbf{u}^k = \mathbf{u}, \\ \lim_{k \rightarrow \infty} \mathbf{p}^k = \mathbf{p}. \end{cases} \quad (6.15)$$

**THEOREM 6.5.** *Assume  $(\mathbf{u}, \mathbf{p}; \lambda)$  is a saddle-point of  $\mathcal{L}_{\text{vtv}}(\mathbf{v}, \mathbf{q}; \mu)$ . Suppose that the minimization problem (6.7) is roughly solved in each iteration, i.e.,  $L = 1$  in Algorithm 6.2. Then the sequence  $(\mathbf{u}^k, \mathbf{p}^k; \lambda^k)$  generated by Algorithm 6.1 satisfies*

$$\begin{cases} \lim_{k \rightarrow \infty} G_{\text{vtv}}(\mathbf{u}^k, \mathbf{p}^k) = G_{\text{vtv}}(\mathbf{u}, \mathbf{p}), \\ \lim_{k \rightarrow \infty} \|\mathbf{p}^k - \nabla \mathbf{u}^k\|_{\mathbf{Q}} = 0. \end{cases} \quad (6.16)$$

Since  $R_{\text{vtv}}(\mathbf{p})$  is continuous, (6.16) indicates that  $\mathbf{u}^k$  is a minimizing sequence of  $F_{\text{vtv}}$ . If we further have  $\text{Null}(\mathbf{K}) = \{0\}$ , then

$$\begin{cases} \lim_{k \rightarrow \infty} \mathbf{u}^k = \mathbf{u}, \\ \lim_{k \rightarrow \infty} \mathbf{p}^k = \mathbf{p}. \end{cases} \quad (6.17)$$

**6.2. Dual methods for vectorial TV model.** In this sub-section we give CGM and Chambolle's dual methods for vectorial TV restoration model.

We start from the optimality condition of the saddle-point problem (6.6), which reads

$$\partial_{\mathbf{v}} \mathcal{L}_{\text{vtv}}(\mathbf{v}, \mathbf{q}, \mu)|_{(\mathbf{u}, \mathbf{p}; \lambda)} = \alpha \mathbf{K}^*(\mathbf{K}\mathbf{u} - \mathbf{f}) + \text{div} \lambda + r \text{div}(\mathbf{p} - \nabla \mathbf{u}) = 0, \quad (6.18)$$

$$\partial_{\mathbf{q}} \mathcal{L}_{\text{vtv}}(\mathbf{v}, \mathbf{q}, \mu)|_{(\mathbf{u}, \mathbf{p}; \lambda)} = \partial R_{\text{vtv}}(\mathbf{p}) + \lambda + r(\mathbf{p} - \nabla \mathbf{u}) \ni 0, \quad (6.19)$$

$$\partial_{\mu} \mathcal{L}_{\text{vtv}}(\mathbf{v}, \mathbf{q}, \mu)|_{(\mathbf{u}, \mathbf{p}; \lambda)} = \mathbf{p} - \nabla \mathbf{u} = 0, \quad (6.20)$$

where  $\partial R_{\text{vtv}}(\mathbf{p})$  is the sub-differential of  $R_{\text{vtv}}$  at  $\mathbf{p}$ , and  $\text{div} \lambda = (\text{div} \lambda_1, \text{div} \lambda_2, \dots, \text{div} \lambda_M)$  as well as  $\text{div}(\mathbf{p} - \nabla \mathbf{u})$  means similarly.

**6.2.1. The CGM dual method.** The CGM dual method for color image restoration is still missing in the literature, as pointed out in [4] that “the question of extension is open for the CGM's model”. Here we present the method via simplifying the optimality condition of the saddle-point problem (6.6). Using (6.20) to eliminate  $\mathbf{p}$  and rearranging the result yield

$$\begin{aligned} \alpha \mathbf{K}^*(\mathbf{K}\mathbf{u} - \mathbf{f}) + \text{div} \lambda &= 0 \\ \nabla \mathbf{u} + \lambda |\nabla \mathbf{u}| &= 0 \end{aligned} \quad (6.21)$$

which is a similar system of that in the CGM dual method in [12]. Newton's linearization techniques can then be used to simultaneously compute the primal and dual variables  $\mathbf{u}$  and  $\lambda$  in (6.21).

**6.2.2. Chambolle's dual method.** If we go a step further, we will get a method similar with Chambolle's [11]. See [4]. From the first equation of (6.21) we have a relation between the primal variable  $\mathbf{u}$  and the dual variable  $\lambda$  as (assume here  $\text{Null}(\mathbf{K}) = \{0\}$ )

$$\mathbf{u} = (\alpha \mathbf{K}^* \mathbf{K})^{-1} (\alpha \mathbf{K}^* \mathbf{f} - \text{div} \lambda). \quad (6.22)$$

Substituting this equation into the second equation of (6.21) gives

$$\nabla((\mathbf{K}^* \mathbf{K})^{-1} (\alpha \mathbf{K}^* \mathbf{f} - \text{div} \lambda)) + \lambda |\nabla((\mathbf{K}^* \mathbf{K})^{-1} (\alpha \mathbf{K}^* \mathbf{f} - \text{div} \lambda))| = 0, \quad (6.23)$$

which can be solved with a semi-implicit gradient descent scheme. Here we derive Chambolle's dual method for vectorial TV model in a different way from [4].

---

**Algorithm 6.3** Split Bregman iteration for vectorial TV model
 

---

1. *Initialization*:  $(\mathbf{u}^{-1}, \mathbf{p}^{-1}) = (0, 0)$ ,  $(\mathbf{g}_u^{-1}, \mathbf{g}_p^{-1}) = (0, 0)$ ;
2. For  $k = 0, 1, 2, \dots$ : Compute  $(\mathbf{u}^k, \mathbf{p}^k)$  from

$$(\mathbf{u}^k, \mathbf{p}^k) = \arg \min_{(\mathbf{u}, \mathbf{p})} D_{G_{\text{tv}}}^{(\mathbf{g}_u^{k-1}, \mathbf{g}_p^{k-1})}((\mathbf{u}, \mathbf{p}), (\mathbf{u}^{k-1}, \mathbf{p}^{k-1})) + \frac{1}{2} \|\mathbf{p} - \nabla \mathbf{u}\|_{\mathbf{Q}}^2, \quad (6.24)$$

and update

$$\begin{aligned} \mathbf{g}_u^k &= \mathbf{g}_u^{k-1} - \operatorname{div}(\mathbf{p}^k - \nabla \mathbf{u}^k), \\ \mathbf{g}_p^k &= \mathbf{g}_p^{k-1} - (\mathbf{p}^k - \nabla \mathbf{u}^k). \end{aligned} \quad (6.25)$$


---

**6.3. Split Bregman iteration for vectorial TV model.** Split Bregman iteration for the restoration problem (6.4) is presented in Algorithm 6.3. Therein the minimization problem (6.24) can be solved using Algorithm 6.2 with  $r = 1$ . To our knowledge, Algorithm 6.3 has not been proposed yet.

One can show the equivalence between augmented Lagrangian method (Algorithm 6.1) and split Bregman iteration (Algorithm 6.3), as done in last section for the ROF model.

**7. Extension to high order models.** We can also extend our method and observations to high order models. As well known, the TV restoration models (e.g., ROF and vectorial TV) suffer from staircase effect; see [45, 41, 13, 15, 5] and references therein. To overcome this, high order models have been proposed [10, 13, 46, 27, 29, 16, 24]. Here we take the Lysaker-Lundervold-Tai (LLT) model [27] as an example. Other high order models can be similarly treated. Moreover, we present the model (which is still denoted as ‘‘LLT’’) and method for multi-valued images, for generality.

Since the LLT model is defined using second order derivatives, we need to introduce second order difference operators. Given  $u \in V$ , we define

$$\begin{aligned} (\mathring{D}_{xx}^{-+} u)_{i,j} &:= (\mathring{D}_x^- (\mathring{D}_x^+ u))_{i,j}, \\ (\mathring{D}_{xy}^{++} u)_{i,j} &:= (\mathring{D}_x^+ (\mathring{D}_y^+ u))_{i,j}, \\ (\mathring{D}_{yx}^{++} u)_{i,j} &:= (\mathring{D}_y^+ (\mathring{D}_x^+ u))_{i,j}, \\ (\mathring{D}_{yy}^{-+} u)_{i,j} &:= (\mathring{D}_y^- (\mathring{D}_y^+ u))_{i,j}, \end{aligned}$$

based on the first order difference operators introduced in Section 2. One may verify that, by the definition above,  $(\mathring{D}_{xy}^{++} u)_{i,j} = (\mathring{D}_{yx}^{++} u)_{i,j}$  holds. In this paper we also use some other second order difference operators such as  $\mathring{D}_{xx}^{+-}$ ,  $\mathring{D}_{xy}^{+-}$ ,  $\mathring{D}_{xy}^{-+}$ ,  $\mathring{D}_{xy}^{--}$ . They can be similarly defined and we omit the details. We now denote the discrete Hessian of  $u$  as

$$Hu = \begin{pmatrix} \mathring{D}_{xx}^{-+} u & \mathring{D}_{xy}^{++} u \\ \mathring{D}_{yx}^{++} u & \mathring{D}_{yy}^{-+} u \end{pmatrix} \in Q_2$$

with

$$(Hu)_{i,j} = \begin{pmatrix} (\mathring{D}_{xx}^{-+} u)_{i,j} & (\mathring{D}_{xy}^{++} u)_{i,j} \\ (\mathring{D}_{yx}^{++} u)_{i,j} & (\mathring{D}_{yy}^{-+} u)_{i,j} \end{pmatrix},$$

where

$$Q_2 = V \times V \times V \times V.$$

We point out that there are actually  $3 \times 4 \times 3$  symmetric discrete Hessians by different combinations of all the second order difference operators. Here we use just one of them defined above. For a vector-valued image  $\mathbf{u} = (u_1, u_2, \dots, u_M) \in \mathbf{V}$ , the Hessian is computed channel by channel, and denoted as

$$H\mathbf{u} = (Hu_1, Hu_2, \dots, Hu_M) \in \mathbf{Q}_2,$$

where

$$\mathbf{Q}_2 = \underbrace{Q_2 \times \dots \times Q_2}_M.$$

Given

$$\mathbf{p} = \left( \left( \begin{array}{cc} p_1^{11} & p_1^{12} \\ p_1^{21} & p_1^{22} \end{array} \right), \left( \begin{array}{cc} p_2^{11} & p_2^{12} \\ p_2^{21} & p_2^{22} \end{array} \right), \dots, \left( \begin{array}{cc} p_M^{11} & p_M^{12} \\ p_M^{21} & p_M^{22} \end{array} \right) \right) \in \mathbf{Q}_2$$

and

$$\mathbf{q} = \left( \left( \begin{array}{cc} q_1^{11} & q_1^{12} \\ q_1^{21} & q_1^{22} \end{array} \right), \left( \begin{array}{cc} q_2^{11} & q_2^{12} \\ q_2^{21} & q_2^{22} \end{array} \right), \dots, \left( \begin{array}{cc} q_M^{11} & q_M^{12} \\ q_M^{21} & q_M^{22} \end{array} \right) \right) \in \mathbf{Q}_2,$$

the inner product and norm in the space  $\mathbf{Q}_2$  are as follows

$$\begin{aligned} (\mathbf{p}, \mathbf{q})_{\mathbf{Q}_2} &= \sum_{1 \leq m \leq M} ((p_m^{11}, q_m^{11})_V + (p_m^{12}, q_m^{12})_V + (p_m^{21}, q_m^{21})_V + (p_m^{22}, q_m^{22})_V), \\ \|\mathbf{p}\|_{\mathbf{Q}_2} &= \sqrt{(\mathbf{p}, \mathbf{p})_{\mathbf{Q}_2}}. \end{aligned}$$

Similarly with those in the ROF and vectorial TV models, we mention the following pixel-by-pixel norm

$$|\mathbf{p}_{i,j}| = \sqrt{\sum_{1 \leq m \leq M} ((p_m^{11})_{i,j}^2 + (p_m^{12})_{i,j}^2 + (p_m^{21})_{i,j}^2 + (p_m^{22})_{i,j}^2)}.$$

By regarding the Hessian as an operator  $H : \mathbf{V} \rightarrow \mathbf{Q}_2$ , we find its adjoint operator  $H^* : \mathbf{Q}_2 \rightarrow \mathbf{V}$  as

$$H^*(\mathbf{p}) = (H^*p_1, H^*p_2, \dots, H^*p_M),$$

where

$$H^*p_m = \mathring{D}_{xx}^{+-} p_m^{11} + \mathring{D}_{yx}^{--} p_m^{12} + \mathring{D}_{xy}^{--} p_m^{21} + \mathring{D}_{yy}^{+-} p_m^{22}.$$

We then consider the following image restoration problem

$$\min_{\mathbf{u} \in \mathbf{V}} \{F_{\text{lit}}(\mathbf{u}) = R_{\text{lit}}(H\mathbf{u}) + \frac{\alpha}{2} \|\mathbf{K}\mathbf{u} - \mathbf{f}\|_{\mathbf{V}}^2\}, \quad (7.1)$$

where

$$R_{\text{lit}}(H\mathbf{u}) = \sum_{1 \leq i,j \leq N} |(H\mathbf{u})_{i,j}|,$$

and  $\mathbf{f} \in \mathbf{V}$  is the observed image, and  $\mathbf{K} : \mathbf{V} \rightarrow \mathbf{V}$  is the blur kernel. Under the following assumption

- $\text{Null}(H) \cap \text{Null}(\mathbf{K}) = \{0\}$ ,

the functional  $F_{\text{lit}}(\mathbf{u})$  in (7.1) is convex, proper, coercive, and continuous. According to the generalized Weierstrass theorem and Fermat's theorem, the problem (7.1) has at least one solution  $\mathbf{u}$ , which is characterized by

$$0 \in \alpha \mathbf{K}^*(\mathbf{K}\mathbf{u} - \mathbf{f}) + H^* \partial R_{\text{lit}}(H\mathbf{u}), \quad (7.2)$$

where  $\partial R_{\text{lit}}(H\mathbf{u})$  is the sub-differential of  $R_{\text{lit}}$  at  $H\mathbf{u}$ . Moreover, if  $\text{Null}(\mathbf{K}) = \{0\}$ , the minimizer is unique.

In the following we present augmented Lagrangian method to solve (7.1). We only give the algorithm. Convergence results and connections to Chambolle's dual method [37] are similar with those in previous sections. It is also quite straightforward to get the other two new methods, i.e., the CGM dual method and split Bregman iteration for this problem, by following our observations in the previous section. Here we omit these details.

We first reformulate (7.1) to be the following constrained optimization problem

$$\begin{aligned} \min_{\mathbf{u} \in \mathbf{V}, \mathbf{p} \in \mathbf{Q}_2} \{ & G_{\text{lit}}(\mathbf{u}, \mathbf{p}) = R_{\text{lit}}(\mathbf{p}) + \frac{\alpha}{2} \|\mathbf{K}\mathbf{u} - \mathbf{f}\|_{\mathbf{V}}^2 \}, \\ \text{s.t.} \quad & \mathbf{p} = H\mathbf{u}. \end{aligned} \quad (7.3)$$

To solve (7.3), we define the augmented Lagrangian functional as

$$\mathcal{L}_{\text{lit}}(\mathbf{v}, \mathbf{q}; \mu) = R_{\text{lit}}(\mathbf{q}) + \frac{\alpha}{2} \|\mathbf{K}\mathbf{v} - \mathbf{f}\|_{\mathbf{V}}^2 + (\mu, \mathbf{q} - H\mathbf{v})_{\mathbf{Q}_2} + \frac{r}{2} \|\mathbf{q} - H\mathbf{v}\|_{\mathbf{Q}_2}^2, \quad (7.4)$$

where  $\mu \in \mathbf{Q}_2$ , and consider the following saddle-point problem:

$$\begin{aligned} \text{Find } & (\mathbf{u}, \mathbf{p}; \lambda) \in \mathbf{V} \times \mathbf{Q}_2 \times \mathbf{Q}_2, \\ \text{s.t. } & \mathcal{L}_{\text{lit}}(\mathbf{u}, \mathbf{p}; \mu) \leq \mathcal{L}_{\text{lit}}(\mathbf{u}, \mathbf{p}; \lambda) \leq \mathcal{L}_{\text{lit}}(\mathbf{v}, \mathbf{q}; \lambda), \forall (\mathbf{v}, \mathbf{q}; \mu) \in \mathbf{V} \times \mathbf{Q}_2 \times \mathbf{Q}_2. \end{aligned} \quad (7.5)$$

Similarly with Theorem 4.1,  $\mathbf{u} \in \mathbf{V}$  is a solution of (7.1) if and only if there exist  $\mathbf{p} \in \mathbf{Q}_2$  and  $\lambda \in \mathbf{Q}_2$  such that  $(\mathbf{u}, \mathbf{p}; \lambda)$  is a solution of (7.5). This can be shown by noticing (7.2).

---

**Algorithm 7.1** Augmented Lagrangian method for the LLT model

---

1. *Initialization:*  $\lambda^0 = 0$ ;
2. For  $k = 0, 1, 2, \dots$ : Compute  $(\mathbf{u}^k, \mathbf{p}^k)$  from

$$(\mathbf{u}^k, \mathbf{p}^k) \approx \arg \min_{(\mathbf{v}, \mathbf{q}) \in (\mathbf{V}, \mathbf{Q}_2)} \mathcal{L}_{\text{lit}}(\mathbf{v}, \mathbf{q}; \lambda^k), \quad (7.6)$$

and update

$$\lambda^{k+1} = \lambda^k + r(\mathbf{p}^k - H\mathbf{u}^k). \quad (7.7)$$


---

An iterative algorithm is given in Algorithm 7.1 to solve the saddle-point problem (7.5). To solve the minimization problem (7.6), we separate it to be the following two sub-problems:

$$\min_{\mathbf{v}} \frac{\alpha}{2} \|\mathbf{K}\mathbf{v} - \mathbf{f}\|_{\mathbf{V}}^2 - (\lambda^k, H\mathbf{v})_{\mathbf{Q}_2} + \frac{r}{2} \|\mathbf{q} - H\mathbf{v}\|_{\mathbf{Q}_2}^2, \quad (7.8)$$

for a given  $\mathbf{q}$ , and

$$\min_{\mathbf{q}} R_{\text{llt}}(\mathbf{q}) + (\lambda^k, \mathbf{q})_{\mathbf{Q}_2} + \frac{r}{2} \|\mathbf{q} - H\mathbf{v}\|_{\mathbf{Q}_2}^2, \quad (7.9)$$

for a given  $\mathbf{v}$ .

Similarly with (6.11), (6.12) and (6.13), we here give the solutions to (7.8) and (7.9) as follows. From the optimality condition of the sub-problem (7.8) and using Fourier transforms, we deduce

$$\alpha \mathcal{F}(\mathbf{K}^*) \mathcal{F}(\mathbf{K}) \mathcal{F}(\mathbf{v}) + r \mathcal{F}(H^*) \mathcal{F}(H) \mathcal{F}(\mathbf{v}) = \alpha \mathcal{F}(\mathbf{K}^*) \mathcal{F}(\mathbf{f}) + \mathcal{F}(H^*) \mathcal{F}(\lambda^k) + r \mathcal{F}(H^*) \mathcal{F}(\mathbf{q}), \quad (7.10)$$

from which  $\mathcal{F}(\mathbf{v})$  and then  $\mathbf{v}$  can be found. The sub-problem (7.9) has the following closed form solution

$$\mathbf{q}_{i,j} = \begin{cases} (1 - \frac{1}{r} \frac{1}{|\mathbf{w}_{i,j}|}) \mathbf{w}_{i,j}, & |\mathbf{w}_{i,j}| > \frac{1}{r}, \\ 0, & |\mathbf{w}_{i,j}| \leq \frac{1}{r}, \end{cases} \quad (7.11)$$

where

$$\mathbf{w} = H\mathbf{v} - \frac{\lambda^k}{r}. \quad (7.12)$$

We then use an iterative procedure to alternatively calculate  $\mathbf{v}$  and  $\mathbf{q}$  according to (7.10) (7.11); see Algorithm 7.2. Here  $L$  can be chosen using some convergence

---

**Algorithm 7.2** Augmented Lagrangian method for the LLT model – solve the minimization problem (7.6)

---

- *Initialization:*  $\mathbf{u}^{k,0} = \mathbf{u}^{k-1}$ ,  $\mathbf{p}^{k,0} = \mathbf{p}^{k-1}$ ;
  - *For*  $l = 0, 1, 2, \dots, L - 1$ : Compute  $\mathbf{u}^{k,l+1}$  from (7.10) for  $\mathbf{q} = \mathbf{p}^{k,l}$ ; and then compute  $\mathbf{p}^{k,l+1}$  from (7.11) for  $\mathbf{v} = \mathbf{u}^{k,l+1}$ ;
  - $\mathbf{u}^k = \mathbf{u}^{k,L}$ ,  $\mathbf{p}^k = \mathbf{p}^{k,L}$ .
- 

test techniques and usually simply set to be  $L = 1$ , as in the ROF and vectorial TV restoration problems.

**8. Examples and discussion.** Several numerical examples are provided in Figures 8.1, 8.2, 8.3, 8.4, and 8.5. In all these figures, SNR and  $t$  denote signal-noise-ratio and the CPU time usage, respectively. In Figures 8.1, 8.2, and 8.3, we show augmented Lagrangian method applied to the ROF restoration model. Examples in Figures 8.4 and 8.5 illustrate the extension of our method to vectorial TV and high order restoration models, respectively. Comparisons between our method and some built-in Matlab functions, i.e. *deconvwnr.m*, *deconvreg.m* and *deconvlucy.m*, are shown in Figures 8.1, 8.2, and 8.4. As one can see, our method generates much better restoration than these built-in Matlab functions in comparable (or even less) CPU time costs. In Figure 8.3, we also compare our method (with increasing parameter  $r$ ) with the recently developed FTVd package based on pure splitting-and-penalty, which is one of the most efficient approaches as compared to other existing methods as discussed in [40]. From Figures 8.1 and 8.3 people can also compare FTVd with our method with fixed parameter  $r$ . The performance of our method is similar with that of the FTVd package. The example illustrated in Figure 8.5 shows augmented Lagrangian method applied to the LLT model to reduce the staircase effect of the ROF model; see the zoomed images in Figure 8.5.

We would like to give some comments on the efficiency of our method. As one can see, our method contains two iterations, one inner iteration and one outer iteration. In the inner iteration (see Algorithm 4.2, 6.2, and 7.2), FFT-based implementation and closed form solution of sub-problems ensure the efficiency. In the outer iteration (see Algorithm 4.1, 6.1, and 7.1, which are equivalent to split Bregman iteration for corresponding problems, e.g., Algorithm 3.1 and 6.3), the method can be interpreted as a gradient ascent approach for the dual variable (the sub-gradients of the regularization term). This is particularly efficient when the regularization term is homogeneous 1, e.g., TV and vectorial TV norms used in this paper.

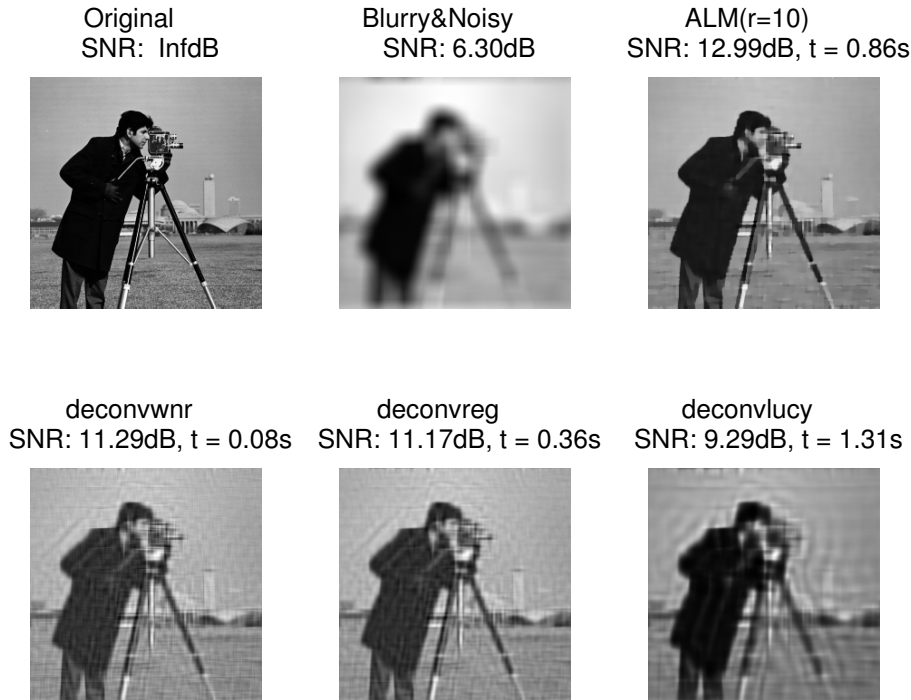


FIG. 8.1. Augmented Lagrangian method (ALM) with parameter  $r = 10$  for ROF restoration, and comparisons to built-in Matlab functions.

**9. Conclusions and future works.** In this paper we present augmented Lagrangian method to solve the ROF model. As demonstrated in the examples, our method benefits from both accuracy and efficiency. We also give some convergence analysis to our approach. Besides, we show close connections between augmented Lagrangian method and several other particularly efficient approaches, such as the CGM and Chambolle's dual methods, as well as split Bregman iteration. In addition, our method and observations are extended to vectorial TV and high order restoration models. In these extensions, one may easily obtain some new methods for vectorial TV and high order models, e.g., the CGM dual method and split Bregman iteration applied to these models. A Possible future work is to further extend the method to models with other data fidelity terms, e.g.,  $TV - L_1$  model. We noticed that recently a variant of the ROF model was proposed in [26] which avoids the staircase effect of the ROF model. To apply our approach to this variant is also valuable for future research.

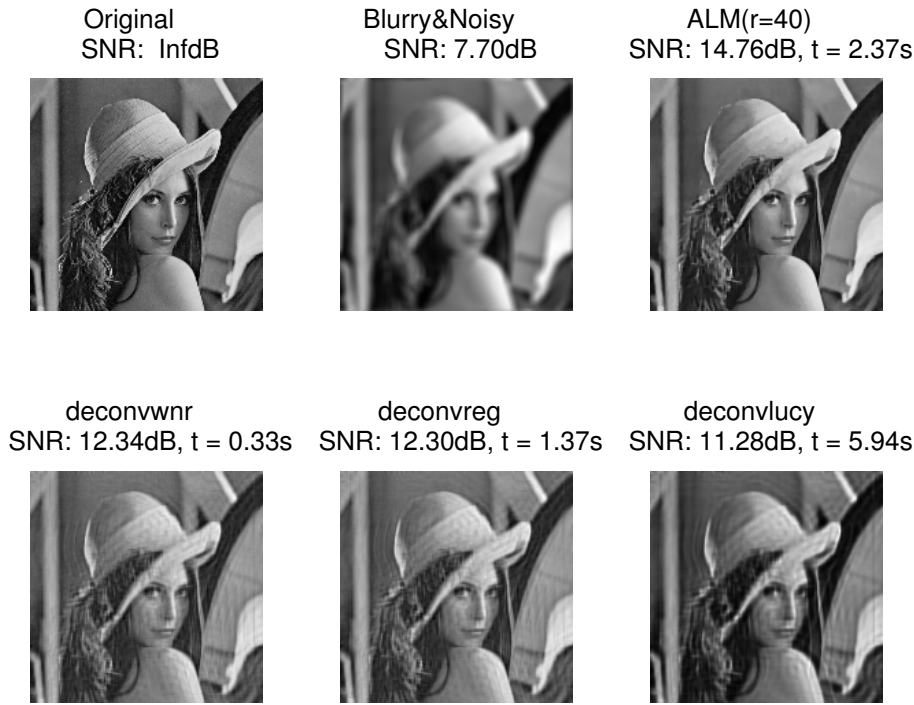


FIG. 8.2. Augmented Lagrangian method (ALM) with parameter  $r = 40$  for ROF restoration, and comparisons to built-in Matlab functions.

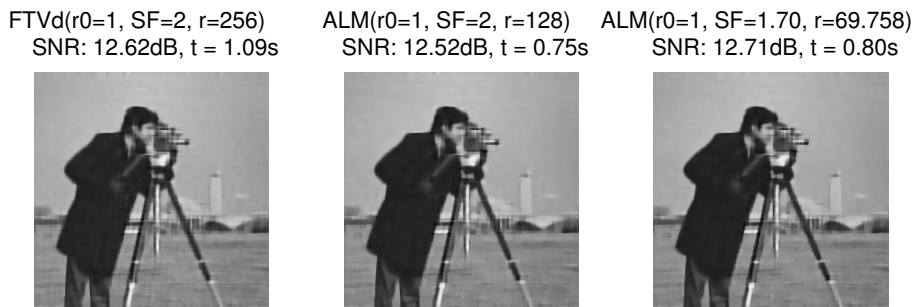


FIG. 8.3. Comparisons between FTVd package (splitting-and-penalty) and augmented Lagrangian method with increasing penalty parameters for ROF restoration. In the sub-figures,  $r_0$ , SF and  $r$  stand for the initial value, the scaling factor and the final value of the penalty parameter of methods, respectively. The blurry&noisy image is shown in Fig. 8.1.

**Acknowledgement.** The research has been supported by MOE (Ministry of Education) Tier II project T207N2202 and IDM project NRF2007IDM-IDM002-010. In addition, support from SUG 20/07 is also gratefully acknowledged.

#### REFERENCES

- [1] D.P. Bertsekas, *Multiplier Methods: A Survey*, Automatica 12, 133–145 (1976).
- [2] P. Blomgren, and T.F. Chan, *Color TV: Total Variation Methods for Restoration of Vector-Valued Images*, IEEE Trans. Image Process., vol. 7, no. 3, pp. 304–309, 1998.





FIG. 8.4. Augmented Lagrangian method (ALM) with parameter  $r = 50$  for vectorial TV restoration, and comparisons to built-in Matlab functions.

- [3] L.M. Bregman, *The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming*, USSR Computational Mathematics and Mathematical Physics, 7(1967), pp. 200–217.
- [4] X. Bresson, and T.F. Chan, *Fast Minimization of the Vectorial Total Variation Norm and Applications to Color Image Processing*, UCLA CAM Report.
- [5] A. Buades, B. Coll, and J.M. Morel, *The Staircasing Effect in Neighborhood Filters and its Solution*, IEEE Trans. Image Process., vol. 15, no. 6, pp. 1499–1505, 2006.
- [6] A. Caboussat, R. Glowinski, and V. Pons, *An Augmented Lagrangian Approach to the Numerical Solution of a Non-Smooth Eigenvalue Problem*, Journal of Numerical Mathematics, to appear.
- [7] J. Cai, S. Osher, and Z. Shen, *Linear Bregman Iterations and Compressed Sensing*, UCLA CAM report 08–06.
- [8] J. Cai, S. Osher, and Z. Shen, *Split Bregman Methods and Frame Based Image Restoration*, UCLA CAM report 09–28.
- [9] J.L. Carter, *Dual Methods for Total Variation - Based Image Restoration*, Ph.D. thesis, UCLA, 2001.
- [10] A. Chambolle, and P.L. Lions, *Image Recovery via Total Variation Minimization and Related Problems*, Numer. Math., 76(1997), pp. 167–188.
- [11] A. Chambolle, *An Algorithm for Total Variation Minimization and Applications*, J. Math. Imaging Vis., 20(2004), pp. 89–97.
- [12] T.F. Chan, G.H. Golub, and P. Mulet, *A Nonlinear Primal-Dual Method for Total Variation-Based Image Restoration*, SIAM J. Sci. Comput., 20(1999), pp. 1964–1977.
- [13] T. Chan, A. Marquina, and P. Mulet, *High-Order Total Variation-Based Image Restoration*, SIAM J. Scientific Comput., 22(2000), pp. 503–516.
- [14] T.F. Chan, S.H. Kang and J.H. Shen, *Total Variation Denoising and Enhancement of Color Images Based on the CB and HSV Color Models*, J. Visual Commun. Image Repres., vol. 12, pp. 422–435, 2001.
- [15] T. Chan, S. Esedoglu, F. Park and A. Yip, *Recent Developments in Total Variation Image Restoration*, CAM report(2005), ftp://ftp.math.ucla.edu/pub/camreport/cam05-01.pdf,

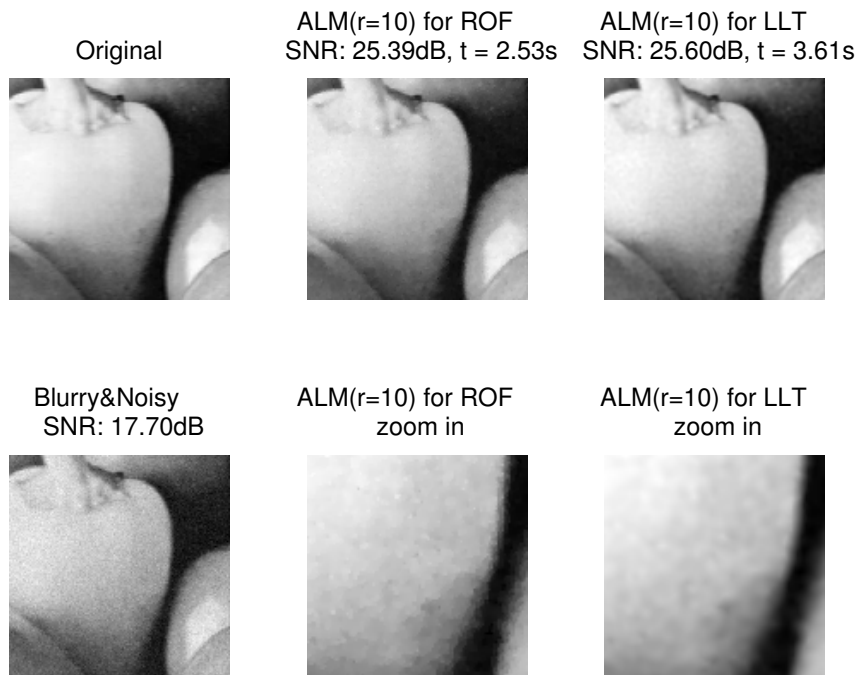


FIG. 8.5. Augmented Lagrangian method (ALM) with parameter  $r = 10$  for ROF and LLT restorations.

- UCLA.
- [16] T.F. Chan, S. Esedoglu, and F.E. Park, *A Fourth Order Dual Method for Staircase Reduction in Texture Extraction and Image Restoration Problems*, UCLA CAM Report, 2005.
  - [17] Tony. F. Chan, K. Chen and Xue-Cheng Tai, *Nonlinear Multilevel Schemes for Solving the Total Variation Image Minimization Problem*, in “Image Processing Based on Partial Differential Equations”, Springer, Heidelberg, 2006.
  - [18] Ke Chen, and Xue-Cheng Tai, *A Nonlinear Multigrid Method for Total Variation Minimization from Image Restoration*, J. Sci. Comput., 33(2007), pp. 115–138.
  - [19] I. Ekeland, and R. Témam, *Convex Analysis and Variational Problems*, SIAM, 1999.
  - [20] E. Esser, *Applications of Lagrangian-Based Alternating Direction Methods and Connections to Split Bregman*, UCLA CAM Report, 09–31.
  - [21] R. Glowinski, P. Le Tallec, *Augmented Lagrangians and Operator-Splitting Methods in Nonlinear Mechanics*, SIAM, Philadelphia (1989).
  - [22] T. Goldstein, and S. Osher, *The Split Bregman Method for L1 Regularized Problems*, SIAM Journal on Imaging Sciences, 2(2009), pp. 323–343.
  - [23] M.R. Hestenes, *Multiplier and Gradient Methods*, Journal of Optimization Theory and Applications 4, 303–320 (1969).
  - [24] W. Hinterberger, and O. Scherzer, *Variational Methods on the Space of Functions of Bounded Hessian for Convexification and Denoising*, Computing, vol. 76, pp. 109–133, 2006.
  - [25] Y. Huang, M. Ng and Y. Wen, *A Fast Total Variation Minimization Method for Image Restoration*, SIAM Multiscale Modeling and Simulation, accepted, 2008.
  - [26] K. Jalalzai, and A. Chambolle, *Enhancement of Blurred and Noisy Images Based on an Original Variant of the Total Variation*, Scale Space and Variational Methods in Computer Vision, Second International Conference, SSVM 2009, Voss, Norway, June 1-5, 2009. Proceedings. Lecture Notes in Computer Science 5567, pp. 368–376, Springer, 2009.

- [27] M. Lysaker, A. Lundervold, and X.-C. Tai, *Noise removal using fourth-order partial differential equation with applications to medical Magnetic Resonance Images in space and time*, IEEE Trans. Image Process., 12(2003), pp. 1579–1590.
- [28] M. Lysaker, S. Osher, and X.-C. Tai, *Noise Removal Using Smoothed Normals and Surface Fitting*, IEEE Trans. Image Process., vol. 13, no. 10, 2004, pp. 1345–1357.
- [29] M. Lysaker, and X.-C. Tai, *Iterative Image Restoration Combining Total Variation Minimization and a Second Order Functional*, Int'l J. Computer Vision, 2005.
- [30] S. Osher, M. Burger, D. Goldfarb, J.J. Xu, W.T. Yin, *An Iterative Regularization Method for Total Variation-Based Image Restoration*, SIAM Multiscale Model. Simul., 4(2005), pp. 460–489.
- [31] M.J.D. Powell, *A Method for Nonlinear Constraints in Minimization Problems*, Optimization, Fletcher, R. ed., Academic Press, New York, 283–298 (1972).
- [32] R.T. Rockafellar, *A Dual Approach to Solving Nonlinear Programming Problems by Unconstrained Optimization*, Mathematical Programming 5, 354–373 (1973).
- [33] L. Rudin, S. Osher, and E. Fatemi, *Nonlinear Total Variation Based Noise Removal Algorithms*, Physica D, 60(1992), pp. 259–268.
- [34] G. Sapiro, and D.L. Ringach, *Anisotropic Diffusion of Multivalued Images with Applications to Color Filtering*, IEEE Trans. Image Process., vol. 5, no. 11, pp. 1582–1586, 1996.
- [35] O. Scherer, *Denoising With Higher Order Derivatives of Bounded Variation and an Application to Parameter Estimation*, Computing, vol. 60, pp. 1–27, 1998.
- [36] S. Setzer, *Split Bregman Algorithm, Douglas-Rachford Splitting and Frame Shrinkage*, Scale Space and Variational Methods in Computer Vision, Second International Conference, SSVM 2009, Voss, Norway, June 1-5, 2009. Proceedings. Lecture Notes in Computer Science 5567, pp. 464-476, Springer, 2009.
- [37] G. Steidl, *A Note on the Dual Treatment of Higher-Order Regularization Functionals*, Computing, vol. 76, pp. 135–148, 2006.
- [38] Xue-Cheng Tai, Chunlin Wu, *Augmented Lagrangian method, dual methods and split Bregman iteration for ROF model*, Scale Space and Variational Methods in Computer Vision, Second International Conference, SSVM 2009, Voss, Norway, June 1-5, 2009. Proceedings. Lecture Notes in Computer Science 5567, pp. 502-513, Springer, 2009.
- [39] Y.L. Wang, W.T. Yin, and Y. Zhang, *A Fast Algorithm for Image Deblurring with Total Variation Regularization*, UCLA CAM Report.
- [40] Y. Wang, J. Yang, W. Yin, and Y. Zhang, *A New Alternating Minimization Algorithm for Total Variation Image Reconstruction*, SIAM Journal on Imaging Sciences, 1(2008), pp. 248–272.
- [41] J. Weickert, *Anisotropic Diffusion in Image Processing*, Stuttgart, B.G. Teubner, 1998.
- [42] J.F. Yang, W.T. Yin, Y. Zhang, and Y.L. Wang, *A Fast Algorithm for Edge-Preserving Variational Multichannel Image Restoration*, UCLA CAM Report 08-50.
- [43] W.T. Yin, S. Osher, D. Goldfarb and J. Darbon, *Bregman Iterative Algorithms for Compressend Sensing and Related Problems*, SIAM J. Imaging Sciences, 1(2008), pp. 143–168.
- [44] W.T. Yin, *Analysis and Generalizations of the Linearized Bregman Method*, UCLA CAM Report 09-42.
- [45] Y.-L. You, W. Xu, A. Tannenbaum, and M. Kaveh, *Behavioral Analysis of Anisotropic Diffusion in Image Processing*, IEEE Trans. Image Process., vol. 5, pp. 1539–1553, 1996.
- [46] Y.-L. You and M. Kaveh, *Fourth-Order Partial Differential Equation for Noise Removal*, IEEE Trans. Image Process., 9(2000), pp. 1723–1730.
- [47] X.Q. Zhang, M. Burger, X. Bresson, and S. Osher, *Bregmanized Nonlocal Regularization for Deconvolution and Sparse Reconstruction*, UCLA CAM Report 09-03, 2009.
- [48] M. Zhu, S.J. Wright, T.F. Chan, *Duality-Based Algorithms for Total Variation Image Restoration*, UCLA CAM Report 08-33, 2008.