Signal Processing 91 (2011) 1933–1940

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Fast optimization for multichannel total variation minimization with non-quadratic fidelity

Juyong Zhang [a,*], Chunlin Wu [b]

[a] Division of Computer Communications, School of Computer Engineering, Nanyang Technological University, Singapore
[b] Department of Mathematics, National University of Singapore, Singapore

ARTICLE INFO

Article history:
Received 4 December 2010
Received in revised form
16 February 2011
Accepted 19 February 2011
Available online 24 February 2011

Keywords:
Image restoration
Impulse noise
Poisson noise
Total variation
Augmented Lagrangian method

ABSTRACT

Total variation (TV) has been proved very successful in image processing, and it has been combined with various non-quadratic fidelities for non-Gaussian noise removal. However, these models are hard to solve because TV is non-differentiable and nonlinear, and non-quadratic fidelity term is also nonlinear and even non-differentiable for some special cases. This prevents their widespread use in practical applications. Very recently, it was found that the augmented Lagrangian method is extremely efficient for this kind of models. However, only the single-channel case (e.g., gray images) is considered. In this paper, we propose a general computational framework based on augmented Lagrangian method for multichannel TV minimization with non-quadratic fidelity, and then show how to apply it to two special cases: L^1 and Kullback–Leibler (KL) fidelities, two common and important data terms for blurry images corrupted by impulsive noise or Poisson noise, respectively. For these typical fidelities, we show that the sub-problems either can be fast solved by FFT or have closed form solutions. The experiments demonstrate that our algorithm can fast restore high quality images.

1. Introduction

In image processing and signal processing areas, many problems can be casted as an energy minimization formulation, whose objective energy functional often contains two terms, one is the regularization term modeling some a priori information about the original data, and a fidelity term measuring some types of deviation of the original data from the observed data. One of the most used and successful regularization term is total variation, which was first put forward in [1], and has been successfully applied in many image processing areas, such as image restoration [1–6], image segmentation, image compression, image reconstruction and surface reconstruction [5,7,8]. Its success is based on the fact that the gradient is sparse in most images and total variation catches this property, like the basis pursuit problem [9] in compressive sensing [10,11].

As for the fidelity term, different problems use different formulation to suit its applications. For image restoration, the most common and popular fidelity term is the L^2 type, which has shown great success for recovering blurry images corrupted by Gaussian noise [1,12,2,13]. However, for many other important applications, the L^2 type fidelity term is not suitable any more and we need to use a non-quadratic fidelity term to depict the relation between the observed data and the original data. For example, it has been shown that we can get high restoration results by using TV-L^1 model to recover images corrupted by blur and impulsive noise [14,3]. Another typical example is the TV-KL model for Poisson noise [4].

Although TV regularization has many good properties, it is also well-known to be hard to solve because of its non-differentiability and nonlinear properties. A lot of effort has been contributed to design fast solvers [15–22] in the past few years, however, all of these works only consider TV

* Corresponding author.
E-mail address: juyong@pmail.ntu.edu.sg (J. Zhang).

0165-1684/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.sigpro.2011.02.015

minimization with squared L^2 fidelity term. Compared with TV-L^2 model, TV with non-quadratic fidelity models are more complicated because the first order variations of these fidelities are no longer linear. Several algorithms have been put forward to overcome these difficulties. For example, for TV-L^1 model, there are gradient descent method [23], LAD method [24], splitting-and-penalty based method [25], and primal-dual method [26] based on semi-smooth Newton algorithm [27], alternating direction methods [28], as well as augmented Lagrangian method [29] for gray images. The same problem exists for TV-KL model, which is effective to deblur images corrupted by Poisson noise, but is quite hard to minimize. Popular methods to solve this TV-KL model are gradient descent [4], multilevel method [30], the scaled gradient projection method [31], and EM-TV alternative minimization [32], as well as variable splitting and convex optimization based methods [33,34].

In this paper, we put forward a general computational framework for multichannel TV (MTV) minimization with non-quadratic fidelity. Our algorithm is derived from the well-known variable-splitting and augmented Lagrangian method in optimization, which is an unconstrained function based on augmenting the Lagrangian function with a quadratic penalty term. Another benefit of this formulation is that two of three subproblems either can be fast solved by FFT or has closed form solution for general fidelity terms, and the remaining subproblem has closed-form formula for some typical fidelities. We will see that our algorithm dramatically improves the computation speed because of the FFT implementation and closed form solutions. Besides, our different parameters for different auxiliary variables strategy is also much more effective than the one in [28,33,34]. Except the MTV regularization, our algorithm can also be extended to effectively handle weighted MTV, as well as high-order regularization terms, like the strategy used in [35]. In general, we present a very efficient algorithm for a quite general type of models which is composed of MTV-like regularization term and non-quadratic fidelity term. As shown in Section 4, our algorithm makes the edge-preserving variational color image restoration models to be practically viable technologies.

The rest of the paper is organized as follows. In the next section, we will first give an introduction about multichannel total variation. Augmented Lagrangian method based numerical solver for multichannel total variation with non-quadratic fidelity terms will be given in Section 3, and its application to two special cases: blurry images corrupted by impulsive noise or Poisson noise will also be presented in this section. Experimental results are shown in Section 4 and we conclude the paper in Section 5.

2. Multichannel total variation

Let us denote in general an M-channel image by $\mathbf{u}=(u_1,u_2,\ldots,u_M)^T$, where $u_m(\forall m=1,2,\ldots,M)$ can be treated as a single channel image represented as an $N\times N$ matrix. For convenience of description, we introduce the following notations:

$$\mathbf{V}=\underbrace{V\times V\times\cdots\times V}_{M},\quad \mathbf{Q}=\underbrace{Q\times Q\times\cdots\times Q}_{M},$$

where V denotes the Euclidean space $\mathbb{R}^{N\times N}$ and $Q=V\times V$. The definition of inner products, norms, and operators ∇ and div in the spaces V, Q, $\mathbf{V}$ and $\mathbf{Q}$ can be referred to [35]. Assume $\mathbf{f}\in\mathbf{V}$ is an observed multichannel image containing both blur and noise. Then it can be in general modeled as

$$\mathbf{u}\xrightarrow{\text{blur}}\mathbf{K}\mathbf{u}\xrightarrow{\text{noise}}\mathbf{f}, \tag{1}$$

where $\mathbf{u}\in\mathbf{V}$ is the true image, $\mathbf{K}:\mathbf{V}\to\mathbf{V}$ is a blur operator. It has the form of

$$\mathbf{K}=\begin{pmatrix} K_{11} & K_{12} & \cdots & K_{1M}\\ K_{21} & K_{22} & \cdots & K_{2M}\\ \vdots & \vdots & \ddots & \vdots\\ K_{M1} & K_{M2} & \cdots & K_{MM}\end{pmatrix},$$

where $K_{ij}\in\mathbb{R}^{N^2\times N^2}$ is a convolution matrix. The diagonal elements of $\mathbf{K}$ denote within channel blur whereas the off diagonal elements describe cross channel blur. Here we do not specify the noise model $\mathbf{n}$, which can be Gaussian, impulsive, Poisson and even others.

Image restoration aims at recovering $\mathbf{u}$ from $\mathbf{f}$. However, recovering $\mathbf{u}$ from $\mathbf{f}$ by directly inverting (1) is an ill-posed problem because small perturbations in the data may produce unbounded variations in the solution. To stabilize the recovery of $\mathbf{u}$, regularization on the solution should be considered, which reflects some a priori preferences. One of the most basic and successful image restoration models is based on the TV regularization, which was first put forward in [1] for single channel images. For multichannel signal $\mathbf{u}$, it is important to couple channels in regularization. For example, TV on single channel image has been extended to "color TV" in [2] and multichannel TV in [12,36,13]. Some other approaches of coupling channels in regularization are the Beltrami flow [37] and the Mumford–Shah functional [38]. In this work, we use multichannel TV:

$$R_{mtv}(\nabla\mathbf{u})=\mathrm{TV}(\mathbf{u})=\sum_{1\le i,j\le N}\sqrt{\sum_{1\le m\le M}|(\nabla u_m)_{i,j}|^2}, \tag{2}$$

which is convenient to compute and produces high quality results.

The multichannel total variation regularization term $R_{mtv}(\nabla\mathbf{u})$ only models some a priori information about $\mathbf{u}$, which is not enough. We need to use a fidelity term $F(\mathbf{Ku},\mathbf{f})$ to measures some types of deviation of $\mathbf{u}$ from $\mathbf{f}$. And thus the functional to be minimized takes the form:

$$\min_{\mathbf{u}\in\mathbf{V}}\{E(\mathbf{u})=R_{mtv}(\nabla\mathbf{u})+F(\mathbf{Ku},\mathbf{f})\}, \tag{3}$$

where the fidelity term depends on the statistic of the noise model. For example, as we discussed in the Introduction, it is a squared L^2 term for the Gaussian noise image restoration, while L^1 term for impulsive noise and KL term for Poisson noise.

Since the blur is essentially averaging, it is reasonable to assume

- Assumption 1. $\mathrm{null}(\nabla)\cap\mathrm{null}(\mathbf{K})=\{0\}$,

where $\mathrm{null}(\cdot)$ represents the null space of $\cdot$.

For the fidelity term, we make the following assumptions:

- Assumption 2. $\mathrm{dom}(R\circ\nabla)\cap\mathrm{dom}(F\circ\mathbf{K})\neq\emptyset$;
- Assumption 3. $F(\mathbf{z})$ is convex, proper, and coercive [39,40];
- Assumption 4. $F(\mathbf{z})$ is continuous over $\mathrm{dom}(F)$,

where $\mathrm{dom}(F)=\{\mathbf{z}\in\mathbf{V}:-\infty<F(\mathbf{z})<+\infty\}$ is the domain of F, with similar definitions for $\mathrm{dom}(R\circ\nabla)$ and $\mathrm{dom}(F\circ\mathbf{K})$.

Under the above assumptions, it can be verified that the functional $E(\mathbf{u})$ in (3) is convex, proper, coercive, and lower semi-continuous. According to the generalized Weierstrass theorem and Fermat's rule [39,40], we can further prove that the problem (3) has at least one solution $\mathbf{u}^*$. Actually, these assumptions are quite general and many fidelities such as L^1, L^2 and KL meet all of them.

3. Augmented Lagrangian method for MTV with non-quadratic fidelity

For MTV with non-quadratic fidelity models, it is quite hard to directly solve it because MTV and fidelity are nonlinear. Similar with augmented Lagrangian method for TV-L^2 model in [35], we use variable splitting technology to change it to minimization problem with equality constraints first. By introducing two new variables $\mathbf{p}\in\mathbf{Q}$ and $\mathbf{z}\in\mathbf{V}$, we reformulate problem (3) to the following equality constrained optimization problem:

$$\begin{aligned}&\min_{\mathbf{u}\in\mathbf{V},\mathbf{p}\in\mathbf{Q},\mathbf{z}\in\mathbf{V}}\{E(\mathbf{u},\mathbf{p},\mathbf{z})=R_{mtv}(\mathbf{p})+F(\mathbf{z},\mathbf{f})\}\\ &\text{s.t.}\quad \mathbf{p}=\nabla\mathbf{u},\\ &\qquad \mathbf{z}=\mathbf{K}\mathbf{u}.\end{aligned}\tag{4}$$

To solve (4), we define the following augmented Lagrangian functional:

$$\begin{aligned}\mathcal{L}(\mathbf{u},\mathbf{p},\mathbf{z};\lambda_{\mathbf{p}},\lambda_{\mathbf{z}})=&R_{mtv}(\mathbf{p})+F(\mathbf{z},\mathbf{f})+(\lambda_{\mathbf{p}},\mathbf{p}-\nabla\mathbf{u})+(\lambda_{\mathbf{z}},\mathbf{z}-\mathbf{K}\mathbf{u})\\&+\frac{r_p}{2}\|\mathbf{p}-\nabla\mathbf{u}\|^2+\frac{r_z}{2}\|\mathbf{z}-\mathbf{K}\mathbf{u}\|^2,\end{aligned}\tag{5}$$

with Lagrange multipliers $\lambda_{\mathbf{p}}\in\mathbf{Q},\lambda_{\mathbf{z}}\in\mathbf{V}$ and positive constants r_p, r_z. We then consider the following saddle-point problem:

$$\begin{aligned}&\text{Find}\quad(\mathbf{u}^*,\mathbf{p}^*,\mathbf{z}^*;\lambda_{\mathbf{p}}^*,\lambda_{\mathbf{z}}^*)\in\mathbf{V}\times\mathbf{Q}\times\mathbf{V}\times\mathbf{Q}\times\mathbf{V}\\ &\text{s.t.}\quad\begin{aligned}&\mathcal{L}(\mathbf{u}^*,\mathbf{p}^*,\mathbf{z}^*;\lambda_{\mathbf{p}},\lambda_{\mathbf{z}})\\&\le\mathcal{L}(\mathbf{u}^*,\mathbf{p}^*,\mathbf{z}^*;\lambda_{\mathbf{p}}^*,\lambda_{\mathbf{z}}^*)\\&\le\mathcal{L}(\mathbf{u},\mathbf{p},\mathbf{z};\lambda_{\mathbf{p}}^*,\lambda_{\mathbf{z}}^*),\\&\forall(\mathbf{u},\mathbf{p},\mathbf{z};\lambda_{\mathbf{p}},\lambda_{\mathbf{z}})\in\mathbf{V}\times\mathbf{Q}\times\mathbf{V}\times\mathbf{Q}\times\mathbf{V}.\end{aligned}\end{aligned}\tag{6}$$

According to optimization theories [39,40], we can verify the following theorem, just like [29].

Theorem 1. *$\mathbf{u}^*\in\mathbf{V}$ is a solution of (3) if and only if there exist $(\mathbf{p}^*,\mathbf{z}^*)\in\mathbf{Q}\times\mathbf{V}$ and $(\lambda_{\mathbf{p}}^*,\lambda_{\mathbf{z}}^*)\in\mathbf{Q}\times\mathbf{V}$ such that $(\mathbf{u}^*,\mathbf{p}^*,\mathbf{z}^*;\lambda_{\mathbf{p}}^*,\lambda_{\mathbf{z}}^*)$ is a solution of (6).*

It guarantees that the solution of the saddle-point problem (6) provides a solution of the original problem (3). In the following Algorithm 1, we present an iterative algorithm to solve the saddle-point problem (6) and address three sub-problems raised up in each iteration.

Algorithm 1. Augmented Lagrangian method for MTV minimization with non-quadratic fidelity.

1. *Initialization*: $\lambda_{\mathbf{p}}^0=0,\lambda_{\mathbf{z}}^0=0$;
2. *For* $k=0,1,2,\ldots$:
   (a) compute $(\mathbf{u}^k,\mathbf{p}^k,\mathbf{z}^k)$ as
   $$(\mathbf{u}^k,\mathbf{p}^k,\mathbf{z}^k)\approx\arg\min_{(\mathbf{u},\mathbf{p},\mathbf{z})\in\mathbf{V}\times\mathbf{Q}\times\mathbf{V}}\mathcal{L}(\mathbf{u},\mathbf{p},\mathbf{z};\lambda_{\mathbf{p}}^k,\lambda_{\mathbf{z}}^k),\tag{7}$$
   where $\mathcal{L}(\mathbf{u},\mathbf{p},\mathbf{z};\lambda_{\mathbf{p}}^k,\lambda_{\mathbf{z}}^k)$ is as in (5);
   (b) update
   $$\lambda_{\mathbf{p}}^{k+1}=\lambda_{\mathbf{p}}^k+r_p(\mathbf{p}^k-\nabla\mathbf{u}^k),$$
   $$\lambda_{\mathbf{z}}^{k+1}=\lambda_{\mathbf{z}}^k+r_z(\mathbf{z}^k-\mathbf{K}\mathbf{u}^k).$$

Since the variables $\mathbf{u},\mathbf{p},\mathbf{z}$ in $\mathcal{L}(\mathbf{u},\mathbf{p},\mathbf{z};\lambda_{\mathbf{p}}^k,\lambda_{\mathbf{z}}^k)$ are coupled together in the minimization problem (7), it is difficult to solve them simultaneously. Therefore we separate the problem to be three sub-problems and apply an alternative minimization. The three sub-problems are as follows:

- **u**-sub problem: Given $\mathbf{p},\mathbf{z}$,
$$\min_{\mathbf{u}\in\mathbf{V}}\left\{(\lambda_{\mathbf{p}}^k,-\nabla\mathbf{u})+(\lambda_{\mathbf{z}}^k,-\mathbf{K}\mathbf{u})+\frac{r_p}{2}\|\mathbf{p}-\nabla\mathbf{u}\|^2+\frac{r_z}{2}\|\mathbf{z}-\mathbf{K}\mathbf{u}\|^2\right\}.\tag{8}$$

- **p**-sub problem: Given $\mathbf{u},\mathbf{z}$,
$$\min_{\mathbf{p}\in\mathbf{Q}}\left\{R_{mtv}(\mathbf{p})+(\lambda_{\mathbf{p}}^k,\mathbf{p})+\frac{r_p}{2}\|\mathbf{p}-\nabla\mathbf{u}\|^2\right\}.\tag{9}$$

- **z**-sub problem: Given $\mathbf{u},\mathbf{p}$,
$$\min_{\mathbf{z}\in\mathbf{V}}\left\{F(\mathbf{z},\mathbf{f})+(\lambda_{\mathbf{z}}^k,\mathbf{z})+\frac{r_z}{2}\|\mathbf{z}-\mathbf{K}\mathbf{u}\|^2\right\}.\tag{10}$$

Note here we omit the constant terms in the objective functionals in (8)–(10).

In the following we will show how to efficiently solve these sub-problems and then present an alternative minimization algorithm to solve (7).

3.1. Solving the **u**-sub problem (8)

Eq. (8) is a quadratic optimization problem, whose optimality condition reads

$$r_z\mathbf{K}^*\mathbf{K}\mathbf{u}-r_p\Delta\mathbf{u}=\mathbf{K}^*\lambda_{\mathbf{z}}^k+r_z\mathbf{K}^*\mathbf{z}-\mathrm{div}\lambda_{\mathbf{p}}^k-r_p\mathrm{div}\mathbf{p},\tag{11}$$

where $\mathbf{K}^*$ is the L_2 adjoint of $\mathbf{K}$. Following [41,18,42,25,22,35,29], we use Fourier transform (and hence FFT implementation) to solve the above linear equation by considering the periodic boundary conditions. Applying Fourier transform to both sides of (11), we have

$$\mathbf{A}\begin{pmatrix}\mathcal{F}(u_1)\\\mathcal{F}(u_2)\\\vdots\\\mathcal{F}(u_M)\end{pmatrix}=\mathcal{F}(\mathbf{K}^*)\begin{pmatrix}\mathcal{F}(\lambda_{z1}^k+r_zz_1)\\\mathcal{F}(\lambda_{z2}^k+r_zz_2)\\\vdots\\\mathcal{F}(\lambda_{zM}^k+r_zz_M)\end{pmatrix}-\mathcal{F}(\mathrm{div})\begin{pmatrix}\mathcal{F}(\lambda_{p1}^k+r_pp_1)\\\mathcal{F}(\lambda_{p2}^k+r_pp_2)\\\vdots\\\mathcal{F}(\lambda_{pM}^k+r_pp_M)\end{pmatrix},\tag{12}$$

where $\mathcal{F}(\mathbf{u})=(\mathcal{F}(u_1),\mathcal{F}(u_2),\ldots,\mathcal{F}(u_M))^T$, $\mathcal{F}(\mathbf{K})=(\mathcal{F}(K_{ij}))_{M\times M}$, $\mathcal{F}(\mathbf{K}^*)=conj(\mathcal{F}(K_{ij}))$ and $\mathbf{A}=(r_z\mathcal{F}(\mathbf{K}^*)\mathcal{F}(\mathbf{K})-r_p\mathcal{F}(\Delta))$. $conj(x)$

here is the complex conjugate of x. Clearly, the matrices $\mathcal{F}(\mathbf{K})$, $\mathcal{F}(\mathbf{K}^*)$, $\mathcal{F}(\Delta)$, $\mathcal{F}(\mathrm{div})$ and thus $\mathbf{A}$ only need to be computed once before iteration.

The $\mathbf{u}$-sub problem can be solved in three steps. First, we apply discrete FFTs to both sides of (11). Then, we solve the resulting systems (12) by Gaussian elimination for $\mathcal{F}(\mathbf{u})$. At last, we apply $\mathcal{F}^{-1}$ to $\mathcal{F}(\mathbf{u})$ to obtain a new $\mathbf{u}$. The total number of $N^2 \times N^2$ size discrete Fourier transforms (including inverse Fourier transforms) is $2M$ per each $\mathbf{u}$-sub problem (8).

If we assume the Neumann boundary conditions and all the blurring kernels are symmetric, the forward and inverse FFTs can be replaced by the forward and inverse discrete cosine transforms (DCTs); see [43]. In this case, a longer CPU time is needed for solving (12) because DCT is generally 3–4 times slower than FFT in MATLAB. In our experiments, we assumed the periodic boundary conditions and thus used FFTs.

3.2. Solving the $\mathbf{p}$-sub problem (9)

According to the definition of $R_{mtv}(\mathbf{p})$ and $\|\cdot\|_{\mathbf{Q}}$, problem (9) equivalents to solving the following problem:

$$\min_{\mathbf{p}\in\mathbf{Q}}\left\{\sum_{1\le i,j\le N}|\mathbf{p}_{i,j}|+\frac{r_p}{2}\sum_{1\le i,j\le N}\left|\mathbf{p}_{i,j}-\left(\nabla\mathbf{u}-\frac{\lambda_{\mathbf{p}}^k}{r_p}\right)_{i,j}\right|^2\right\}.$$

As one can see, it is decomposable and the problem takes the following form at each pixel (i,j):

$$\min_{\mathbf{q}\in\mathbb{R}^{2M}}\left\{|\mathbf{q}|+\frac{r_p}{2}|\mathbf{q}-\mathbf{w}|^2\right\}, \tag{13}$$

where $\mathbf{w}\in\mathbb{R}^{2M}$.

Similarly with [41,18,44,22,35,29], (9) has the following closed form solution:

$$\mathbf{p}_{i,j}=\max\left(0,1-\frac{1}{r_p|\mathbf{w}_{i,j}|}\right)\mathbf{w}_{i,j}, \tag{14}$$

where

$$\mathbf{w}=\nabla\mathbf{u}-\frac{\lambda_{\mathbf{p}}^k}{r_p}\in\mathbf{Q}.$$

3.3. Solving the $\mathbf{z}$-sub problem (10)

For a general fidelity $F(\cdot)$, there is no reason to find a closed form solution for (10). Fortunately, the objective functional in (10) is strictly convex, proper, coercive and lower semi continuous. Therefore, (10) has a unique solution and can be obtained by various numerical optimization methods. At the same time, for some special and typical (non-quadratic) fidelities, we still have closed form solutions; see the two special cases in Section 3.4. Our method is therefore particularly efficient for these typical and important fidelities.

After knowing how to solve (8)–(10), we now present the following alternative minimization procedure (Algorithm 2) to solve (7).

Algorithm 2. Augmented Lagrangian method for TV restoration with non-quadratic fidelity—solve the minimization problem (7).

- *Initialization*: $\mathbf{u}^{k,0}=\mathbf{u}^{k-1},\mathbf{p}^{k,0}=\mathbf{p}^{k-1},\mathbf{z}^{k,0}=\mathbf{z}^{k-1}$;
- *For* $l=0,1,2,\ldots,L-1$:
  - compute $\mathbf{u}^{k,l+1}$ from (12) for $\mathbf{p}=\mathbf{p}^{k,l},\mathbf{z}=\mathbf{z}^{k,l}$;
  - compute $\mathbf{p}^{k,l+1}$ from (14) for $\mathbf{u}=\mathbf{u}^{k,l+1}$, $\mathbf{z}=\mathbf{z}^{k,l}$;
  - compute $\mathbf{z}^{k,l+1}$ by solving (10) for $\mathbf{u}=\mathbf{u}^{k,l+1},\mathbf{p}=\mathbf{p}^{k,l+1}$;
- $\mathbf{u}^k=\mathbf{u}^{k,L},\mathbf{p}^k=\mathbf{p}^{k,L},\mathbf{z}^k=\mathbf{z}^{k,L}$.

According to our numerical tests, we found that $L=1$ setting is a good choice in Algorithm 2, which also coincides with the conclusion in [21,29].

3.4. Applications

We have put forward a quite general computational framework based on augmented Lagrangian method for MTV minimization with non-quadratic fidelity, which is a quite common model in signal processing. In this part, we just show two special cases: to restore blurry images corrupted by impulsive noise or Poisson noise.

For image restoration, the most used is the TV-L^2 model, which is quite effective to restore blurry images corrupted by Gaussian noise. However, the noises in many data may not obey Gaussian distribution. For example, the noises in malfunctioning pixels in camera sensors, faulty memory locations in hardware and erroneous transmission always belong to impulsive noise. Poisson noise is another very common noise, which is often contained in signals in various applications such as radiography, fluorescence microscopy, positron-emission-tomography (PET), optical nanoscopy and astronomical imaging applications [4,32]. For these two typical noises, it has been shown that L^1 or KL fidelity is quite effective to restore the degraded images [14,3,4,25,32]. In this section, by applying Algorithms 1 and 2 to MTV-L^1 restoration for recovering blurry images corrupted by impulsive noise (e.g., salt-and-pepper noise and random-valued noise), and MTV-KL restoration for recovering blurry images corrupted by Poisson noise, we will see that these two models can be efficiently minimized because their $\mathbf{z}$-sub problems (10) can also be fast solved by closed form solution except the $\mathbf{u}$-sub problem and $\mathbf{p}$-sub problem.

3.4.1. Augmented lagrangian method for MTV-L^1 restoration

As introduced in the above, MTV-L^1 model is especially effective to restore blurry images corrupted by impulsive noise. It aims at solving the following minimization problem:

$$\min_{\mathbf{u}\in\mathbf{V}}\{E_{MTV-L^1}(\mathbf{u})=R_{mtv}(\nabla\mathbf{u})+\alpha\|\mathbf{K}\mathbf{u}-\mathbf{f}\|_{L^1}\}, \tag{15}$$

which is a special case of (3) where the fidelity term is

$$F(\mathbf{K}\mathbf{u},\mathbf{f})=\alpha\|\mathbf{K}\mathbf{u}-\mathbf{f}\|_{L^1}.$$

We apply Algorithms 1 and 2 to solve (15). For this special fidelity, $\mathbf{p}$-sub problem and $\mathbf{z}$-sub problem of (15) have similar formulations. Using the same way as in Section 3.2, we have the following explicit solution for

the **z**-sub problem (10):

$$\mathbf{z}_{i,j} = \mathbf{f}_{i,j} + \max\left(0, 1 - \frac{\alpha}{r_z|\mathbf{w}_{i,j} - \mathbf{f}_{i,j}|}\right)(\mathbf{w}_{i,j} - \mathbf{f}_{i,j}), \tag{16}$$

where

$$\mathbf{w} = \mathbf{K}\mathbf{u} - \frac{\lambda_{\mathbf{z}}^k}{r_z} \in \mathbf{V}.$$

Fig. 1. Recovered from blur with salt-and-pepper noise from 30% to 60%. The blur operator is the one in (19). All the results are obtained by using the MTV-L^1 model (15) with different numerical solvers (image size: 512×512).

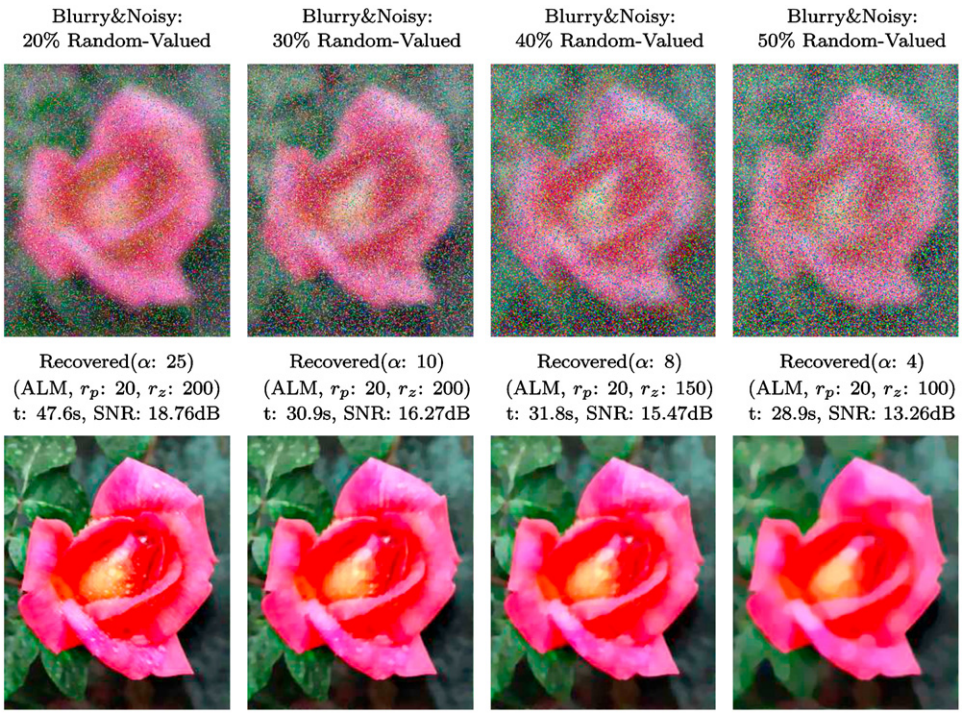

Fig. 2. Recovered from blur with random-valued noise from 20% to 50%. The blur operator is the one in (19). All the results are obtained by using the MTV-L^1 model (15) with different numerical solvers (image size: 250×303).

3.4.2. Augmented Lagrangian method for MTV-KL restoration

To restore blurry images corrupted by Poisson noise, KL divergence is used as the data fidelity, which has an explanation in the context of Bayesian statistics [4]. Similarly with the derivation in [4] for single channel image, under the assumption that the values of $(f_m)_{i,j}$ at the channel m for pixel (i,j) are independent, we deduce the MTV-KL variational model for multichannel images as follows:

$$\min_{\mathbf{u}\in\mathbf{V}}\Big\{E_{MTV-KL}(\mathbf{u})=R_{mtv}(\nabla\mathbf{u})+\alpha\sum_{1\le i,j\le N1}\sum_{1\le m\le M}(((Ku)_m)_{i,j} -(f_m)_{i,j}\log((Ku)_m)_{i,j}) : ((Ku)_m)_{i,j}>0,\forall(m,i,j)\Big\}. \tag{17}$$

Problem (17) is also a special case of (3) where

$$F(\mathbf{Ku},\mathbf{f})=\begin{cases}\alpha\sum\limits_{1\le i,j\le N1}\sum\limits_{1\le m\le M}(((Ku)_m)_{i,j}-(f_m)_{i,j}\log((Ku)_m)_{i,j}) & ((Ku)_m)_{i,j}>0,\\ +\infty, & \text{otherwise.}\end{cases}$$

Therefore, Algorithms 1 and 2 can be applied to compute (17). For this special fidelity, we also have, by considering $(z_m)_{i,j}>0$, a closed form solution to the $\mathbf{z}$-sub problem (10):

$$(z_m)_{i,j}=\frac{1}{2}\left(\sqrt{\left((w_m)_{i,j}-\frac{\alpha}{r_z}\right)^2+4\frac{\alpha}{r_z}(f_m)_{i,j}}+\left((w_m)_{i,j}-\frac{\alpha}{r_z}\right)\right), \tag{18}$$

where

$$\mathbf{w}=\mathbf{Ku}-\frac{\lambda_{\mathbf{z}}^k}{r_z}\in\mathbf{V}.$$

4. Experimental results

Our proposed augmented Lagrangian algorithm is a quite general framework suitable for multichannel TV regularization with non-quadratic fidelity. In this section, we test our algorithm on recovering blurry color images with impulsive noise or Poisson noise on several levels. In the rest of this section, we will first describe restoration results for impulsive noise and then for Poisson noise.

The algorithms are implemented in MATLAB and all blurring effects are generated using the MATLAB function "imfilter33" with periodic boundary conditions. We first blurred the image by cross-channel blurring described below and then corrupted its pixels by noise. let $A(hsize)$ denotes the average blur of the size *hsize*,$G(hsize, sigma)$ the Gaussian blur of the size *hsize* and standard deviation *sigma*, and $M(len, theta)$ the motion blur with motion length *len* and angle *theta*. We first define

$$H(1)=A(13),\quad H(2)=A(15),\quad H(3)=A(17),$$
$$H(4)=G(11,9)\quad H(5)=G(21,11)\quad H(6)=G(31,13),$$
$$H(7)=M(21,45)\quad H(8)=M(41,90)\quad H(9)=M(61,135),$$

and then by making a random permutation to $H(i)$, $i\in\{1,2,\ldots,9\}$, we get H_{ij} $(i,j\in\{1,2,3\})$. Considering that within-channel blurs are usually stronger than cross-channel ones, we assigned larger weights to the within-channel blurs. Similar methods for choosing kernel weights are used in the literature; see, e.g., [25]. Then we get the following blur kernel:

$$\begin{pmatrix}0.7\cdot H_{11} & 0.15\cdot H_{12} & 0.15\cdot H_{13}\\ 0.1\cdot H_{21} & 0.8\cdot H_{22} & 0.1\cdot H_{23}\\ 0.2\cdot H_{31} & 0.2\cdot H_{32} & 0.6\cdot H_{33}\end{pmatrix}. \tag{19}$$

In the experiments, we find that the types locations and kernel size appear to have little influence on the efficiency of our algorithm.

As is usually done, the quality of restoration is measured by the signal-to-noise ration (SNR)

$$SNR\triangleq 10*\log_{10}\frac{\|\overline{\mathbf{u}}-\mathbf{E}(\overline{\mathbf{u}})\|^2}{\|\overline{\mathbf{u}}-\mathbf{u}\|^2},$$

where $\mathbf{E}(\overline{\mathbf{u}})$ is the mean intensity value of the original image $\overline{\mathbf{u}}$, and $\mathbf{u}$ is the restored image.

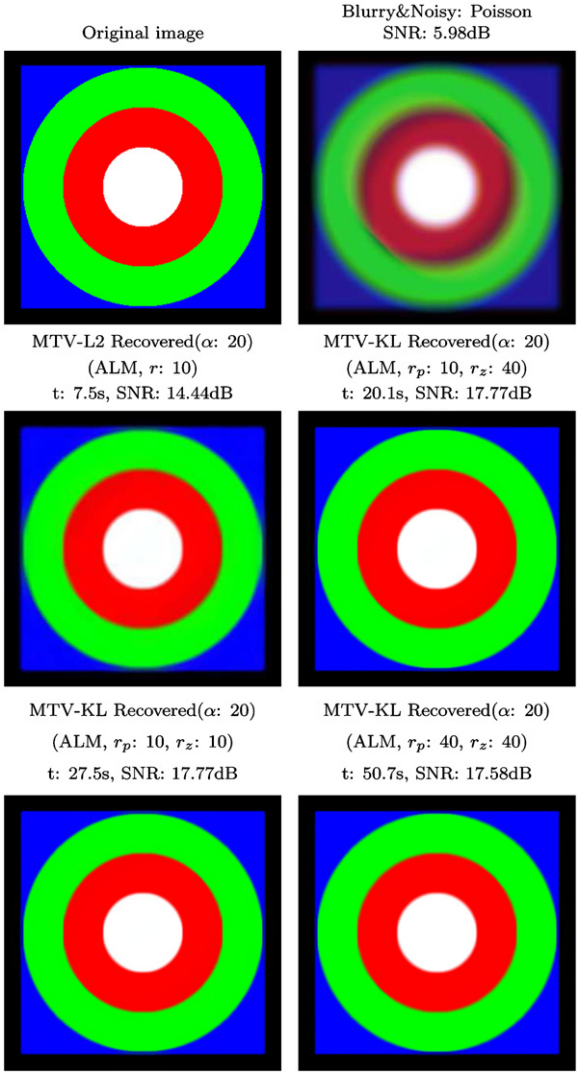

Fig. 3. MTV-L^2 and MTV-KL restoration: recovering from blurry image with Poisson noise. MTV-KL produces much better result than MTV-L^2. Also a better computational efficiency with comparable restoration result (similar SNRs) is achieved by allowing $r_p\neq r_z$ in MTV-KL restoration.

In Figs. 1 and 2, the images are first blurred by the kernel given in (19) and then corrupted by different level of salt-and-pepper noise or random-valued noise. We use MTV-L^1 model to restore them. In each figure, α, t, and SNR denote the parameters of the model, the CPU cost (in seconds), and the SNR of the image, respectively. Note here we use the same α's for all the methods in each example, since our goal is to compare the efficiency of different methods for the same model. All of our experiments were performed under Windows XP and MATLAB R2007a running on a PC with Intel Core 2.66 GHz CPU and 2 GB RAM.

The parameter r_p in our ALM algorithm for salt-and-pepper noise is set to 10 and r_z is around 200; while r_p for random-valued noise is set to 20 and r_z is about 150. As can be seen from Figs. 1 and 2, we can fast get high quality restoration results by using our augmented Lagrangian algorithm. The potential reason for this advantage may be as follows. First, in our method, we simply set L=1 for inner iteration and hence do not need to compute those residuals for stopping criterion, which are calculated in [25]. Second, augmented Lagrangian method benefits from its Lagrange multipliers update, which can be actually interpreted as sub-gradients update in split Bregman iteration [21], and makes the method extremely efficient for homogeneous 1 objective functionals.

As explained in Section 3.4, our algorithm is also very efficient to solve MTV-KL model, which is quite effective to restore blurry image corrupted by Poisson noise. In Fig. 3, we compare the restoration results of MTV-L^2 [35] and MTV-KL models calculated by augmented Lagrangian method (ALM) with parameters r and r_p,r_z, respectively. We find that r_p = 10, r_z = 40 setting generates good results for most pictures. As one can see, MTV-L^2 removes the noise, but has difficulty to preserve sharp edges, while MTV-KL model produces much better results than MTV-L^2. In addition, MTV-KL model can still be calculated very efficiently by augmented Lagrangian method, which is much faster than existing gradient descent method for TV-KL model [4]. In one word, augmented Lagrangian method for MTV-KL model produces much better results than MTV-L^2 model with an acceptable CPU cost, which is quite useful for recovering blurry images with Poisson noise.

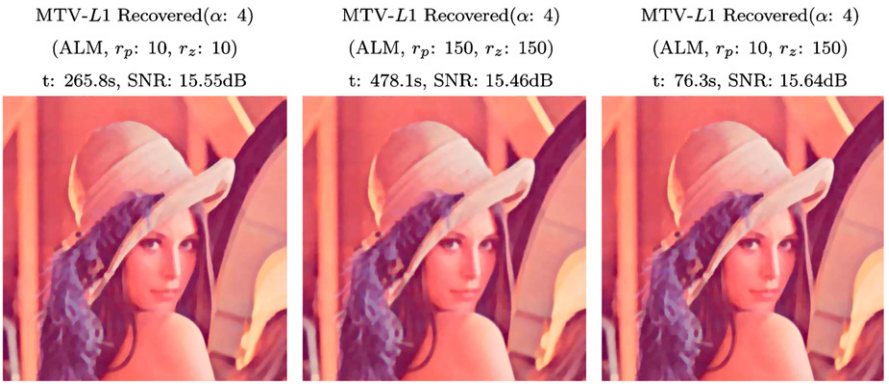

Fig. 4. MTV-L^1 restoration from blurry image with 60% salt-and-pepper noise: a better computational efficiency with comparable restoration result (similar SNRs) is achieved by letting $r_p \neq r_z$.

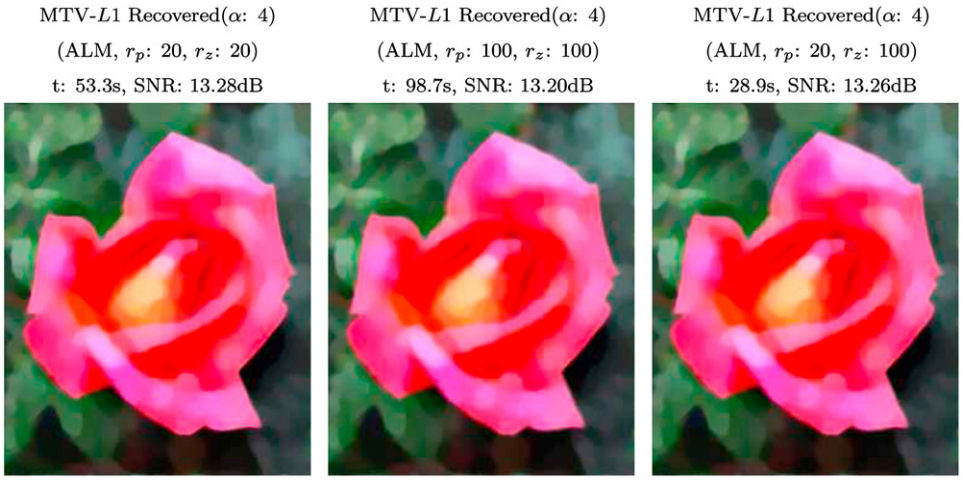

Fig. 5. MTV-L^1 restoration from blurry image with 50% random-valued noise: a better computational efficiency with comparable restoration result (similar SNRs) is achieved by letting $r_p \neq r_z$.

Alternating direction method [28] and variable splitting based methods [33,34] are quite similar with our algorithm, however, they treat the operators in a compact way so that penalty parameters for different auxiliary variables must be the same. In augmented Lagrangian framework, we find that it is more efficient to use different parameters for different auxiliary variables as in our algorithm. From Figs. 3, 4, and 5, we can see that much more efficiency can be achieved by using different penalty parameters, which is quite obvious in the example of Fig. 4.

5. Conclusion

In this paper, we first proposed a fast computational framework for a quite general model, which can be composed of multichannel TV regularization with non-quadratic fidelity term. We then applied it to two typical image deblurring problems with impulsive noise or Poisson noise. Benefitting from FFT implementation, closed form solutions for its sub-problems and simple stopping criterion of the inner iteration, our algorithm is extremely efficient, which is also highly enhanced by the flexibility of $r_p \neq r_z$ in the proposed augmented Lagrangian framework. The experimental results confirm that our proposed algorithm can fast restore high quality results and already achieves a remarkable practical performance.

References

[1] L.I. Rudin, S. Osher, E. Fatemi, Physica D 60 (1992) 259–268.
[2] P. Blomgren, T.F. Chan, IEEE Transactions on Image Processing 7 (1998) 304–309.
[3] M. Nikolova, Journal of Mathematical Imaging and Vision 20 (2004) 99–120.
[4] T. Le, R. Chartrand, T. Asaki, Journal of Mathematical Imaging and Vision 27 (2007) 257–263.
[5] E.J. Candès, F. Guo, Signal Processing 82 (2002) 1519–1543.
[6] J.P. Oliveira, J.M. Bioucas-Dias, M.A.T. Figueiredo, Signal Processing 89 (2009) 1683–1693.
[7] Thomas Pock, Daniel Cremers, Horst Bischof, Antonin Chambolle, An algorithm for minimizing the Mumford-Shah functional, ICCV (2009) 1133–1140.
[8] T. Goldstein, X. Bresson, S. Osher, Journal of Science Computing 45 (2009) 272–293.
[9] S.S. Chen, D.L. Donoho, M.A. Saunders, SIAM Journal on Scientific Computing 20 (1998) 33–61.
[10] D.L. Donoho, IEEE Transactions on Information Theory 52 (2006) 1289–1306.
[11] E.J. Candès, J.K. Romberg, T. Tao, IEEE Transactions on Information Theory 52 (2006) 489–509.
[12] G. Sapiro, D.L. Ringach, IEEE Transactions on Image Processing 5 (1996) 1582–1586.
[13] X. Bresson, T.F. Chan, Inverse Problems and Imaging 2 (2008) 455–484.
[14] M. Nikolova, SIAM Journal on Numerical Analysis 40 (2002) 965–994.
[15] T.F. Chan, G.H. Golub, P. Mulet, SIAM Journal on Scientific Computing 20 (1999) 1964–1977.
[16] A. Chambolle, Journal of Mathematical Imaging and Vision 20 (2004) 89–97.
[17] Y. Huang, M.K. Ng, Y.-W. Wen, Multiscale Modeling and Simulation 7 (2008) 774–795.
[18] Y. Wang, J. Yang, W. Yin, Y. Zhang, SIAM Journal on Imaging Sciences 1 (2008) 248–272.
[19] M. Zhu, T.F. Chan, An efficient primal-dual hybrid gradient algorithm for total variation image restoration, Technical Report 08-34, UCLA CAM, 2008.
[20] M. Zhu, S.J. Wright, T.F. Chan, Duality-based algorithm for total variation image restoration, Technical Report 08-33, UCLA CAM, 2008.
[21] T. Goldstein, S. Osher, SIAM Journal on Imaging Sciences 2 (2009) 323–343.
[22] X.-C. Tai, C. Wu, in: Second International Conference on Scale Space and Variational Methods in Computer Vision, pp. 502–513.
[23] T.F. Chan, S. Esedoglu, SIAM Journal on Applied Mathematics 65 (2005) 1817–1837.
[24] H. Fu, M. Ng, M. Nikolova, J. Barlow, SIAM Journal on Scientific Computing 27 (2006) 1881–1902.
[25] J. Yang, Y. Zhang, W. Yin, SIAM Journal on Scientific Computing 31 (2009) 2842–2865.
[26] Y. Dong, M. Hintermüller, M. Neri, SIAM Journal on Imaging Sciences 2 (2009) 1168–1189.
[27] M. Hintermüller, K. Ito, K. Kunisch, SIAM Journal on Optimization 13 (2002) 865–888.
[28] E. Esser, Applications of Lagrangian-based alternating direction methods and connections to split Bregman, Technical Report 09-31, UCLA CAM, 2009.
[29] Chunlin Wu, Juyong Zhang, Xue-cheng Tai, Augmented Lagrangian method for total variation restoration with non-quadratic fidelity, Inverse Problems and Imaging 5(1) (2011) 237–261.
[30] R.H. Chan, K. Chen, International Journal of Computer Mathematics 84 (2007) 1183–1198.
[31] R. Zanella, P. Boccacci, L. Zanni, M. Bertero, Inverse Problems 25 (2009).
[32] C. Brune, A. Sawatzky, M. Burger, in: Second International Conference on Scale Space and Variational Methods in Computer Vision, pp. 235–246.
[33] M.A.T. Figueiredo, J.M. Bioucas-Dias, in: IEEE Workshop on Statistical Signal Processing, 2009.
[34] S. Setzer, G. Steidl, T. Teuber, Journal of Visual Communication and Image Representation 21 (2010) 193–199.
[35] C. Wu, X.-C. Tai, SIAM Journal on Imaging Sciences 3 (2010) 300–339.
[36] B. Tang, G. Sapiro, V. Caselles, IEEE Transactions on Image Processing 10 (2001) 701–707.
[37] N.A. Sochen, R. Kimmel, R. Malladi, IEEE Transactions on Image Processing 7 (1998) 310–318.
[38] L. Bar, A. Brook, N.A. Sochen, N. Kiryati, IEEE Transactions on Image Processing 16 (2007) 1101–1111.
[39] R. Glowinski, P.L. Tallec, Augmented Lagrangian and Operator Splitting Methods in Nonlinear Mechanics, SIAM, Philadelphia, PA, USA, 1989.
[40] I. Ekeland, R. Téman, Convex Analysis and Variational Problems, SIAM, Philadelphia, PA, USA, 1999.
[41] Y. Wang, W. Yin, Y. Zhang, A fast algorithm for image deblurring with total variation regularization, Technical Report 07-22, UCLA CAM, 2007.
[42] J. Yang, W. Yin, Y. Zhang, Y. Wang, SIAM Journal on Imaging Sciences 2 (2009) 569–592.
[43] M.K. Ng, R.H. Chan, W.-C. Tang, SIAM Journal on Scientific Computing 21 (1999) 851–866.
[44] A. Caboussat, R. Glowinski, V. Pons, Journal of Numerical Mathematics 17 (2009) 3–26.