

# Linearized Proximal Alternating Minimization Algorithm for Motion Deblurring by Nonlocal Regularization

Sangwoon Yun<sup>a</sup>, Hyenkyun Woo<sup>b,\*</sup>

<sup>a</sup>*School of Computational Sciences, Korea Institute for Advanced Study, Seoul 130-722, Korea*

<sup>b</sup>*Institute of Mathematical Sciences, Ewha Womans University, Seoul 120-750, Korea*

---

## Abstract

Non-blind motion deblurring problems are highly ill-posed and so it is quite difficult to find the original sharp and clean image. To handle ill-posedness of the motion deblurring problem, we use nonlocal total variation (abbreviated as TV) regularization approaches. Nonlocal TV can restore periodic textures and local geometric information better than local TV. But, since nonlocal TV requires weighted difference between pixels in the whole image, it demands much more computational resources than local TV. By using the linearization of the fidelity term and the proximal function, our proposed algorithm does not require any inversion of blurring operator and nonlocal operator. Therefore the proposed algorithm is very efficient for motion deblurring problems. We compare the numerical performance of our proposed algorithm with that of several state-of-the-art algorithms for deblurring problems. Our numerical results show that the proposed method is faster and more robust than state-of-the-art algorithms on motion deblurring problems.

*Keywords:* Convex optimization, nonlocal, total variation, motion deblurring, deconvolution, regularization, alternating minimization

---

\*Corresponding author, Tel:+82 2 32776993; Fax:+82 2 32776991.

*Email addresses:* yswcs@kias.re.kr (Sangwoon Yun), hyenkyun@gmail.com (Hyenkyun Woo)

## 1. Introduction

The motion deblurring problem is to recover a sharp and clean image from the given blurred image, which is mainly caused by unsteady movements of a camera [1]. Let  $\Omega \subset \mathbb{R}^2$ ,  $b$  be the given blurred image, and  $k$  be the blurring kernel. Then we wish to find the unknown true image  $u : \Omega \rightarrow \mathbb{R}$  from the given blurred image and the blurring kernel:

$$b = k \otimes u + \eta \tag{1}$$

where  $\int_{\Omega} k(x)dx = 1$  and  $k \geq 0$ ,  $\eta$  is the Gaussian noise, and  $\otimes$  is the convolution operator (with some boundary condition). For simplicity, we assume that the blurring kernel  $k$  is spatially invariant, i.e., a blurred object looks same regardless of its location in the given image. If the spatially invariant blurring kernel is unknown, then the given problem becomes a blind deconvolution. If the given kernel is known, the problem is a non-blind deconvolution. For the blind deconvolution problem, motion blur is easily estimated by using  $\ell_1$  regularization approach [2, 3] because of the strong sparsity of motion blur. For more details on the blind deconvolution, see [1, 2, 3, 4, 5].

In this paper, we only consider the non-blind motion deblurring problems. Even though, convolution/blurring kernel is already known, it is highly ill-posed problem and so it is quite difficult to find the original sharp and clean image. The reason is obvious since the blurring kernel is a kind of a low pass filter and tends to reduce high frequency parts such as textures and edges. Hence directly inverting kernel without using appropriate regularizer causes highly ringing artifacts around edges and textures of an image. To handle this problem, proper regularization methods are required. The most successful regularizer is the local TV [6] used in deblurring problems [4, 7]. The main advantage of using local TV is that it preserves edges due to its linear penalty on differences between adjacent pixels. But, it tends to flatten inhomogeneous areas, such as textures; see Figure 1 (d). To overcome this shortcoming, nonconvex anisotropic TV regularization techniques [1, 5] based on statistical distribution of the gradient of an image or spatially adaptive TV regularization techniques [8, 9] are used. In this paper, we use nonlocal TV regularization technique [10, 11, 12, 13] to restore periodic texture and edge information of the given blurred and noisy image. Nonlocal TV uses the whole image information instead of using adjacent pixel information. In other words, by averaging the current pixel to the other pixels with similar

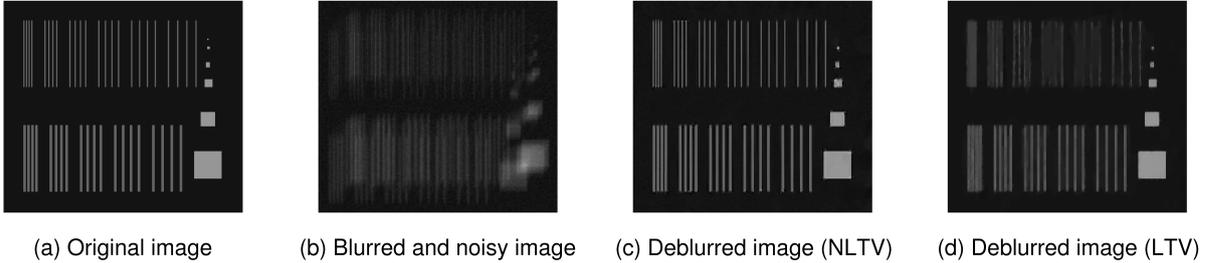


Figure 1: Comparison of the deblurred image by using nonlocal TV regularization with that by using local TV regularization. The texture with periodic patterns and sharp edges in the deblurred image by nonlocal TV are recovered better than that by local TV. We note that the restored images are acquired by our proposed algorithm.

structure neighborhoods, i.e. patches, we can restore the texture with periodic patterns and the sharp edge from the blurred image; see Figure 1 (c). We note that nonlocal total variation is also an efficient approach for other image restoration problems such as denoising, superresolution, compressive sampling, inpainting, and segmentation [14, 15, 16, 17].

Since nonlocal TV requires weighted difference between pixels in the whole image, it consumes more computational resources than local TV. Hence efficient algorithms for solving nonlocal TV deblurring problems are demanded. Recently, Bregmanized operator splitting algorithm [13] has been proposed to solve nonlocal TV deblurring problems. But this method is not quite efficient, since it uses the split Bregman algorithm for solving the nonlocal TV denoising subproblem for each outer iteration and requires the inverses involving a nonlocal Laplacian operator for inner iterations.

Recently developed several other algorithms, such as the alternating minimization algorithm [18, 19] and the primal-dual hybrid gradient algorithm [20], for solving image deblurring problems based on local TV regularization, can be applied to solve nonlocal TV deblurring problems without solving nonlocal TV denoising subproblems, i.e., there are no inner iterations. But we note that those algorithms require to compute the inverses involving a blurring operator [18, 20] or a nonlocal Laplacian operator [19].

The purpose of this paper is to develop a new fast algorithm, for solving

nonlocal TV deblurring problems, which does not require any inverse of the operators. We adapt a similar framework as for the alternating minimization algorithm proposed by Tseng [21], i.e., uses an alternating minimization scheme that used in [18, 19]. As we mentioned above, the current state-of-the-art algorithms, such as the Bregmanized operator splitting algorithm [13], the alternating minimization algorithm [18, 19] and the primal-dual hybrid gradient algorithm [20], require to compute the inverses involving a blurring operator or a nonlocal Laplacian operator. This can take much computational resources. Hence this motivates us to develop new algorithms which use a linearization and proximal techniques to overcome the drawback; see Section 3 for details. Our proposed algorithm also does not require inner iterations which are needed for the Bregmanized operator splitting algorithm. The proposed algorithm is faster and more robust than state-of-the-art algorithms [13, 19, 20] for solving nonlocal TV deblurring problems.

In the motion deblurring problem, since we do not know the boundary condition of the given motion blurred image  $b$  in equation (1), we need to give an appropriate boundary condition on the convolution operator. We consider the reflexive or periodic boundary conditions in this paper. For the periodic boundary condition, we use the Fast Fourier Transform to solve the deblurring problem. The main advantage of FFT is that it only requires  $O(n \log(n))$  arithmetic operations for a convolution of the given image of size  $n$ , regardless of size of a convolution operator [18, 20]. But the periodic boundary condition is too artificial and so it induces strong boundary artifacts in the restored image. To resolve this problem, various techniques are described in [22, 23, 24]. In this paper, we use “edgetaper” function in MATLAB for our numerical experiments with the periodic boundary condition. For the reflexive boundary condition, since the motion blur does not have any specific structure, we can not use any fast transform based method [25].

### 1.1. Nonlocal Total Variation

To solve motion deblurring problems, one may consider the following local TV based variational formulation :

$$\min_u \mu \int_{\Omega} |\nabla u| + \frac{1}{2} \|k \otimes u - b\|_2^2. \quad (2)$$

where  $\mu > 0$ . As we mentioned earlier, fine periodic structures are not well recovered with this model (see Figure 1). To overcome this problem, nonlocal TV regularization techniques have been proposed [14, 16, 17]. In the sequel,

we use the definitions and notations of the nonlocal functionals introduced in [17] to define nonlocal total variation. Let  $w : \Omega \times \Omega \rightarrow \mathbb{R}^+ \cup \{0\}$  be a symmetric weight function, i.e.  $w(x, y) = w(y, x)$ . The nonlocal partial derivative at  $x$  to  $y$  is written as

$$\partial_y u(x) := (u(y) - u(x))\sqrt{w(x, y)}.$$

The nonlocal gradient operator  $\nabla_w u : \Omega \rightarrow \Omega \times \Omega$  is defined as the vector of all partial derivatives at  $x \in \Omega$  :

$$\nabla_w u(x, y) := (u(y) - u(x))\sqrt{w(x, y)}, \quad \text{for all } y \in \Omega,$$

where  $w(x, y)$  is the weight function between  $x$  and  $y$  defined based on the image  $u$ . The nonlocal divergence of a vector  $\varrho : \Omega \times \Omega \rightarrow \mathbb{R}$  at  $x \in \Omega$  can be defined by the adjoint relation with the nonlocal gradient:

$$\langle \nabla_w u, \varrho \rangle = -\langle u, \text{div}_w \varrho \rangle,$$

which defines the nonlocal divergence as

$$\text{div}_w \varrho(x) := \int_{\Omega} (\varrho(x, y) - \varrho(y, x))\sqrt{w(x, y)}dy.$$

Now, we can define nonlocal total variation functional by

$$\begin{aligned} TV_w(u) &= \int_{\Omega} |\nabla_w u| dx \\ &= \int_{\Omega} \sqrt{\int_{\Omega} (u(y) - u(x))^2 w(x, y) dy} dx \end{aligned} \quad (3)$$

Before we proceed, we need to define the weight function  $w(x, y)$ . Let the region  $\Omega_w(x) \subseteq \Omega$  be a neighborhood around  $x \in \Omega$  where the weights are positive. The weight function is defined by the solution of the following energy minimization problem [14, 26]:

$$\min_w E(w) \quad (4)$$

where

$$E(w) = \int_{\Omega} \int_{\Omega_w(x)} [w(x, y) \frac{d_u^2(x, y)}{2h^2} + w(x, y) \log(w(x, y)) - w(x, y)] dy dx, \quad (5)$$

$$d_u^2(x, y) := \int_{\Omega} G_a(t)(F_u(x-t) - F_u(y-t))^2 dt \quad (6)$$

$G_a$  is the Gaussian kernel with standard deviation  $a$ ,  $F_u(z) = u(z) \in B(z)$ , where  $B(z)$  is a patch centered at  $z$ , and  $h$  is the scaling parameter which determines the similarity between different patches. We do not normalize the weight as did in [12, 13]. We refer [10] for different types of metric (6) between  $x$  and  $y$  in  $\Omega$ . The solution of (4) is the following function:

$$w(x, y) = e^{-\frac{d_u^2(x, y)}{2h^2}}, \quad x \in \Omega, y \in \Omega_w(x). \quad (7)$$

To find true image  $u$  in (1), we consider the following formulation:

$$\min_{w, u} F(w, u) := E(w) + \mu TV_w(u) + \frac{1}{2} \|k \otimes u - b\|_2^2, \quad (8)$$

for the motion deblurring model. In this paper, we use an alternating approach that is considered in [14, 26]:

$$\textbf{Weight Update:} \quad w^k = \arg \min_w F(w, u^{k-1}) \quad (9a)$$

$$\textbf{Image Update:} \quad u^k = \arg \min_u F(w^k, u). \quad (9b)$$

We note that Algorithm 1 in Section 4 describes the whole process for non-blind motion deblurring problems in detail.

### 1.2. Notation

For notational convenience, we use vector notation, i.e., the 2D  $M \times N$  image is columnwisely stacked into a vector, for the rest of the paper. Therefore, the unknown true image  $u$  is a vector in  $\mathbb{R}^n$  ( $n = MN$ ), the observed blurred and noisy image  $b$  is a vector in  $\mathbb{R}^n$ , and the motion blur can be modeled as a large sparse  $n \times n$  matrix  $A$  (with an appropriate boundary condition). Then (1) can be expressed as follows:

$$b = Au + \eta, \quad (10)$$

where  $\eta \sim \mathcal{N}(0, \sigma^2)$ . Therefore, in the sequel, we consider the following formulation for the image update part (9b):

$$\min_u \mu \|\nabla_w u\| + \frac{1}{2} \|Au - b\|_2^2, \quad (11)$$

where  $\mu > 0$  and  $\|\nabla_w u\| = \sum_{i=1}^n \|(\nabla_w u)_i\|_2$  with  $(\nabla_w u)_i \in \mathbb{R}^D$ . We note that  $D$  is the size of ‘‘neighbor’’, where the weight (7) is positive.

### 1.3. Overview

The paper is organized as follows. In Section 2, we review several algorithms for solving the TV regularization problem. In Section 3, we describe our proposed algorithm for solving the nonlocal TV deblurring problem. In Section 4, we describe the motion deblurring process and empirically discuss about the permissible range of a stepsize which is crucial for the stability of the proposed algorithm. In Section 5, we report our numerical results on image deblurring problems. In Section 6, we give our conclusions.

## 2. Related Works

The local TV regularization (2) has been popular ever since its introduction by Rudin et al. [6]. Many researchers have proposed algorithms for solving (2) and its variants. Recently, Goldstein and Osher [27] proposed the split Bregman algorithm for solving (2) by using the Bregman iteration to solve the linear constraint reformulation of it:

$$\begin{aligned} \min_{u,z} \quad & \mu \|z\| + \frac{1}{2} \|Au - b\|_2^2, \\ \text{s.t} \quad & \nabla u = z, \end{aligned} \tag{12}$$

and using an alternating approach to approximate the minimization over  $u$  and  $z$ . Esser et al. [28] proposed a modified version of the primal-dual hybrid gradient algorithm proposed by Zhu and Chan [20]. Zhang et al. [29] proposed a unified primal-dual algorithm framework for two classes of problems, that arise in various signal and image processing applications, such as  $\ell_1$  basis pursuit and TV- $\ell_2$  minimization. Wang et al. [18] proposed a fast TV deblurring method that solves a penalty approximation of (2). But, in the penalty approach, a sufficiently large parameter for the penalty term has to be chosen and so it causes numerical difficulties in computation. To overcome this drawback, Tao and Yang [30] proposed an alternating direction method for the problem (2). Independently, Afonso et al. [31] proposed a split augmented Lagrangian shrinkage algorithm, which is an instance of the alternating direction method, for solving one of the standard formulation of image reconstruction that can be formulated as an unconstrained optimization problem, in which the objective function consists of an  $\ell_2$ -fidelity term and a nonsmooth regularizer, that includes the problem (2). Xiao and Yang [19] proposed an inexact alternating minimization algorithm for the problem (2) to avoid evaluating the inverse of the matrix involving  $A^T A$  as well as to prevent from driving the penalty parameter to infinity.

### 2.1. Optimization Algorithms for Total Variation Deblurring Problems

In this subsection, we review three optimization algorithms - the Bregmanized operator splitting with the splitting Bregman (abbreviated as BOSSB) method [13], the primal-dual hybrid gradient (abbreviated as PDHG) algorithm proposed by Zhu and Chan [20], and the inexact alternating direction method (abbreviated as IADM) proposed by Xiao and Yang [19].

Zhang et al. [13] proposed an algorithm to solve the following general equality constrained minimization problem by the Bregman iteration and operator splitting:

$$\begin{aligned} \min_u \quad & J(u) \\ \text{s.t.} \quad & Au = b, \end{aligned} \quad (13)$$

where  $J$  is a general convex functional. Hence the process of the BOSSB applied to solve (13) with the TV, i.e.,  $J(u) = \|\nabla u\|$ , can be expressed as follows:

$$\begin{cases} v^{k+1} = u^k - \delta A^T(Au^k - b^k) \\ u^{k+1} = \arg \min_u \mu \|\nabla u\| + \frac{1}{2\delta} \|u - v^{k+1}\|_2^2 \\ b^{k+1} = b^k + b - Au^{k+1}, \end{cases} \quad (14)$$

where  $\mu > 0$  and  $0 < \delta < 1/\|A^T A\|_2$  with  $\|A^T A\|_2 = \max_{\|x\|_2=1} \|A^T Ax\|_2$ . In what follows, the shrinkage operator is the generalized shrinkage operator defined in [18]:

$$z_i = \mathit{shrink}(a_i, c) = \max(\|a_i\|_2 - c, 0) \frac{a_i}{\|a_i\|_2},$$

where  $a_i \in \mathbb{R}^a$  ( $a$  is 2 for local TV and  $D$  for nonlocal TV) and  $c \in \mathbb{R}$ .

The  $u^{k+1}$  in (14) is founded by the splitting Bregman method [27] and the process can be given as follows ( $u^{k+1} = \lim_{l \rightarrow \infty} \tilde{u}_l$ ):

$$\begin{cases} \tilde{u}_{l+1} = \arg \min_{\tilde{u}} \frac{1}{2\delta} \|\tilde{u} - v^{k+1}\|_2^2 + \frac{\alpha}{2} \|z_l - \nabla \tilde{u} - t_l\|_2^2 \\ z_{l+1} = \arg \min_z \mu \|z\| + \frac{\alpha}{2} \|z - \nabla \tilde{u}_{l+1} - t_l\|_2^2 \\ t_{l+1} = t_l - \nabla \tilde{u}_{l+1} + z_{l+1}, \end{cases} \quad (15)$$

where  $\alpha > 0$ . In a more simplified form, we have

$$\begin{cases} \tilde{u}_{l+1} = (I - \delta\alpha\Delta)^{-1}(v^k - \delta\alpha \operatorname{div}(z_l - t_l)) \\ z_{l+1} = \mathit{shrink}(\nabla \tilde{u}_{l+1} + t_l, \frac{\mu}{\alpha}) \\ t_{l+1} = t_l - \nabla \tilde{u}_{l+1} + z_{l+1}, \end{cases} \quad (16)$$

where  $\Delta$  denotes the Laplacian operator. The BOSSB for solving (13) with  $J(u) = \|\nabla u\|$  requires to compute the inverse of  $I - \delta\alpha\Delta$  at each inner iteration. It consumes more computational resources when the BOSSB is applied to solve the nonlocal TV deblurring problem. In addition, the BOSSB has inner loops as shown in the above. These are main drawbacks when this algorithm is applied to solve the motion deblurring problem.

The PDHG algorithm proposed in [20] targets a saddle point formulation of the problem:

$$\min_u \left( \max_p \langle p, \nabla u \rangle - J^*(p) + \frac{1}{2} \|Au - b\|_2^2 \right),$$

where  $J^*(p) = \begin{cases} 0 & \text{if } \|p\|_* \leq \mu, \\ \infty & \text{otherwise} \end{cases}$ , where  $\|\cdot\|_*$  is a dual norm defined by  $\|x\|_* = \max_{\|y\| \leq 1} \langle x, y \rangle$ , and proceeds by alternating proximal steps that alternately maximize and minimize a penalized form of the saddle function:

$$\begin{cases} p^{k+1} = \arg \max_p -J^*(p) + \langle p, \nabla u^k \rangle - \frac{1}{2\tau} \|p - p^k\|_2^2 \\ u^{k+1} = \arg \min_u \frac{1}{2} \|Au - b\|_2^2 + \langle p^{k+1}, \nabla u \rangle + \frac{1}{2\delta} \|u - u^k\|_2^2, \end{cases} \quad (17)$$

where  $\delta, \tau > 0$ . In a more simplified form, we have

$$\begin{cases} p^{k+1} = \arg \min_{\|p\|_* \leq \mu} - \langle p, \nabla_w u^k \rangle + \frac{1}{2\tau} \|p - p^k\|_2^2 \\ u^{k+1} = (\frac{1}{\delta} I + A^T A)^{-1} (A^T b + \text{div}(p^{k+1}) + \frac{1}{\delta} u^k). \end{cases} \quad (18)$$

The PDHG has attracted much interest in image denoising problem (i.e.,  $A = I$ ) since it outperforms other popular methods such as Chambolle's method [32], the split Bregman [27], and the fast gradient-based algorithm proposed in [33]; see [34] for more details.

As indicated in [28], if  $\delta = \infty$ , then it corresponds to the proximal forward backward splitting method [35] applied to solve the dual formulation of the problem (2) and also corresponds to the alternating minimization algorithm (abbreviated as AMA), which is proposed by Tseng [21], applied to solve the problem (12). The convergence of the general form of the PDHG algorithm has not been proved yet.

Xiao and Yang [19] introduce a fast alternating minimization scheme, for solving (2), which uses a linearization and proximal techniques. In other

words, they introduce an inexact alternating direction method which uses the framework of the alternating direction method of multipliers. But the minimization of the augmented Lagrangian function:

$$\mathcal{L}_\alpha^A(u, z, p) := \frac{1}{2}\|Au - b\|_2^2 + \mu\|z\| + \langle p, z - \nabla u \rangle + \frac{\alpha}{2}\|z - \nabla u\|_2^2, \quad (19)$$

with respect to  $u$ , is solved inexactly in the sense that, by using a linearization of  $\frac{1}{2}\|Au - b\|_2^2$  and adding a proximal term, the approximation of (19) is minimized. The IADM can be expressed as follows:

$$\begin{cases} u^{k+1} = \arg \min_u \langle p^k, z^k - \nabla u \rangle + \langle A^T(Au^k - b), u - u^k \rangle \\ \quad \quad \quad + \frac{\alpha}{2}\|z^k - \nabla u\|_2^2 + \frac{1}{2\delta}\|u - u^k\|_2^2 \\ z^{k+1} = \arg \min_z \mu\|z\| + \langle p^k, z - \nabla u^{k+1} \rangle + \frac{\alpha}{2}\|z - \nabla u^{k+1}\|_2^2 \\ p^{k+1} = p^k + \alpha(z^{k+1} - \nabla u^{k+1}), \end{cases} \quad (20)$$

where  $\alpha > 0$  and  $0 < \delta < 1/\|A^T A\|_2$ . In a more simplified form, we have

$$\begin{cases} u^{k+1} = (\frac{1}{\delta}I - \alpha\Delta)^{-1} ((\frac{1}{\delta}I - A^T A)u^k + A^T b - \text{div}(\alpha z^k + p^k)), \\ z^{k+1} = \text{shrink}(\nabla u^{k+1} - \frac{p^k}{\alpha}, \frac{\mu}{\alpha}), \\ p^{k+1} = p^k + \alpha(z^{k+1} - \nabla u^{k+1}). \end{cases} \quad (21)$$

The global convergence is followed directly by the analysis in [36]. The IADM requires to evaluate the inverse of  $\frac{1}{\delta}I - \alpha\Delta$  at each iteration. This consumes more computational resources when the IADM is applied to solve the nonlocal TV deblurring problem. This is the main drawback when this algorithm is applied to solve the nonlocal TV deblurring problem.

### 3. Linearized Proximal Alternating Minimization Algorithm for Nonlocal Total Variation

In this section, we describe our proposed linearized proximal alternating minimization algorithm (abbreviated as LPAMA) for solving the nonlocal TV deblurring problem of the form (11). In our proposed algorithm, we approximate the  $\ell_2$  fidelity term of the problem (11) by a strongly convex quadratic function. In order to obtain this approximation, we use the linearization of the  $\ell_2$  fidelity term and the proximal function. Then we use a similar framework as for the alternating minimization algorithm (abbreviated as AMA) proposed by Tseng [21] to solve separable convex programming.

First of all, we introduce the AMA for solving separable convex programming. We consider the following separable convex minimization problem

$$\begin{aligned} \min_{u,z} \quad & f(u) + g(z) \\ \text{s.t.} \quad & Mu = z, \end{aligned} \quad (22)$$

where  $f : \mathbb{R}^{n_1} \rightarrow \mathbb{R}$ ,  $g : \mathbb{R}^{n_2} \rightarrow \mathbb{R}$  are convex lower semicontinuous (lsc) functions, and  $M : \mathbb{R}^{n_1} \rightarrow \mathbb{R}^{n_2}$  is an  $n_2 \times n_1$  matrix. We further assume that  $f$  is strongly convex with modulus  $\sigma > 0$ , i.e., for any  $\beta \in (0, 1)$ ,

$$\begin{aligned} & \beta f(u_1) + (1 - \beta)f(u_2) - f(\beta u_1 + (1 - \beta)u_2) \\ & \geq \sigma\beta(1 - \beta)\|u_1 - u_2\|_2^2, \quad \forall u_1, u_2 \in \mathbb{R}^{n_1}, \end{aligned}$$

The Lagrangian function and the augmented Lagrangian function for (22) are respectively

$$\mathcal{L}(u, z, p) := f(u) + g(z) + \langle p, z - Mu \rangle, \quad (23)$$

and

$$\mathcal{L}_\alpha(u, z, p) := f(u) + g(z) + \langle p, z - Mu \rangle + \frac{\alpha}{2}\|z - Mu\|_2^2, \quad (24)$$

Then the AMA for solving (22) can be expressed as the following templates:

$$\begin{cases} u^{k+1} = \arg \min_u \mathcal{L}(u, z^k, p^k) \\ z^{k+1} = \arg \min_z \mathcal{L}_{\alpha^k}(u^{k+1}, z, p^k) \\ p^{k+1} = p^k + \alpha^k(z^{k+1} - Mu^{k+1}), \end{cases} \quad (25)$$

where  $\{\alpha^k\}$  is any sequence of scalars satisfying  $\epsilon \leq \alpha^k \leq 4\sigma/\|M\|_2^2 - \epsilon$  and  $\epsilon$  is any fixed positive scalar not exceeding  $2\sigma/\|M\|_2^2$ . It was shown that this algorithm is convergent if the problem (22) is feasible, the function  $g(z) + \|z\|_2^2$  has minimum, and (22) has an optimal Lagrangian multiplier vector corresponding to the constraints  $Mu = z$ .

When we apply the AMA to solve nonlocal TV denoising problem (i.e. (22) with  $f(u) = \frac{1}{2}\|u - b\|_2^2$ ,  $g(z) = \mu\|z\|$ , and  $Mu = \nabla_w u$ ), we have an advantage of not evaluating any inverses of the operator involving the nonlocal Laplacian operator  $\Delta_w$ . We note that the BOSSB and the IADM require the inverse of the operator involving  $\Delta_w$  even for this denoising problem but the PDHG algorithm does not require the inverse of the operator involving  $\Delta_w$ . Hence the PDHG is quite efficient for the TV denoising problem [34].

If the AMA is directly applied to solve (11), then the first step in the framework (25) is as follows:

$$u^{k+1} = \arg \min_u \mathcal{L}(u, z^k, p^k) = \arg \min_u \mu \|z^k\| + \langle p^k, z^k - \nabla_w u \rangle + \frac{1}{2} \|Au - b\|_2^2 \quad (26)$$

Then

$$u^{k+1} = (A^T A)^{-1} (A^T b - \operatorname{div}_w p^k).$$

But the motion blur, i.e. blurring operator  $A$ , is highly ill-conditioned. This difficulty can be handled by using Tikonov regularization:

$$u^{k+1} = (A^T A + \frac{1}{\hat{\delta}} I)^{-1} (A^T b - \operatorname{div}_w p^k), \quad (27)$$

where  $\hat{\delta} > 0$ . Since the  $u^{k+1}$  in (27) is a unique solution of the following minimization:

$$\min_u \mu \|z^k\| + \langle p^k, z^k - \nabla_w u \rangle + \frac{1}{2} \|Au - b\|_2^2 + \frac{1}{2\hat{\delta}} \|u\|_2^2, \quad (28)$$

we use different fidelity function  $\frac{1}{2} \|Au - b\|_2^2 + \frac{1}{2\hat{\delta}} \|u\|_2^2$  instead of  $\frac{1}{2} \|Au - b\|_2^2$  in this case. Hence the model is different from what we consider in this paper even though we can handle ill-conditionedness of the operator  $A$ . We also note that we need to evaluate the inverse of an operator  $A^T A + \frac{1}{\hat{\delta}} I$ . If we replace the term  $\|u\|_2^2$  by the proximal term  $\|u - u^k\|_2^2$ , i.e.,

$$\min_u \mu \|z^k\| + \langle p^k, z^k - \nabla_w u \rangle + \frac{1}{2} \|Au - b\|_2^2 + \frac{1}{2\hat{\delta}} \|u - u^k\|_2^2 \quad (29)$$

then it becomes the relaxed AMA described in [28] which is equivalent to the PDHG algorithm (18). This relaxed AMA can also handle the ill-conditionedness of the operator  $A$  and the objective function of (29) is a better approximation of the Lagrangian function in (26) than that of (28) when  $\|u^{k+1} - u^k\|_2$  is small. But it also requires evaluating the inverse of an operator  $A^T A + \frac{1}{\hat{\delta}} I$ .

Instead of adding the Tikonov regularization or the proximal function to the equation (26) when we update  $u^{k+1}$ , we consider a different approach. By using Taylor expansion at  $u^k$ , the  $\ell_2$  fidelity term in (26) can be expressed as follows:

$$\frac{1}{2} \|Au - b\|_2^2 = \frac{1}{2} \|Au^k - b\|_2^2 + \langle A^T (Au^k - b), u - u^k \rangle + \frac{1}{2} (u - u^k)^T A^T A (u - u^k).$$

If we replace the quadratic term  $\frac{1}{2}(u - u^k)^T A^T A(u - u^k)$  of the above Taylor expansion by  $\frac{1}{2}(u - u^k)^T H^k(u - u^k)$  with well-conditioned positive definite matrix  $H^k$ . Then the first step in the framework (25) becomes as follows:

$$\begin{aligned} u^{k+1} = \arg \min_u \quad & \mu \|z^k\| + \langle p^k, z^k - \nabla_w u \rangle + \frac{1}{2} \|Au^k - b\|_2^2 \\ & + \langle A^T(Au^k - b), u - u^k \rangle + \frac{1}{2} (u - u^k)^T H^k (u - u^k). \end{aligned}$$

This approach can also resolve the ill-conditionedness of  $A$  and gives a better approximation when  $\|u^{k+1} - u^k\|_2$  is small. The choice of well-posed  $H^k$  is crucial for the performance of this approach. If we let  $H^k = A^T A + \frac{1}{\delta} I$ , then this version becomes the relaxed AMA (29). Less computational resources are demanded at each iteration if we can avoid evaluating the inverses of the operator involving  $A^T A$ . Hence we choose  $H^k$  as a simple constant multiple of the identity matrix, i.e.,  $H_k = \frac{1}{\delta} I$  with  $\delta > 0$ . In other words, we propose to minimize the strongly convex quadratic approximation of (26) by using the linearization of the  $\ell_2$  fidelity term  $\frac{1}{2} \|Au - b\|_2^2$  and adding the proximal function:

$$\arg \min_u \mu \|z^k\| + \langle p^k, z^k - \nabla_w u \rangle + \langle A^T(Au^k - b), u - u^k \rangle + \frac{1}{2\delta} \|u - u^k\|_2^2 \quad (30)$$

instead of minimizing the Lagrangian function (26).

We now describe formally the LPAMA for solving (11) (solving the constraint formulation (12) with nonlocal TV).

**LPAMA:**

Let  $u^0$ ,  $z^0$ , and  $p^0$  be given. Choose  $\delta > 0$ . For  $k = 0, 1, 2, \dots$ , generate  $u^{k+1}$ ,  $z^{k+1}$ ,  $p^{k+1}$  from  $u^k$ ,  $z^k$ ,  $p^k$  according to the following iteration:

**Step 1.** Set  $u^{k+1} = \arg \min_u \langle p^k, z^k - \nabla_w u \rangle + \langle A^T(Au^k - b), u - u^k \rangle + \frac{1}{2\delta} \|u - u^k\|_2^2$ .

**Step 2.** Set  $z^{k+1} = \arg \min_z \mu \|z\| + \langle p^k, z - \nabla_w u^{k+1} \rangle + \frac{\alpha^k}{2} \|z - \nabla_w u^{k+1}\|_2^2$ .

**Step 3.** Set  $p^{k+1} = p^k + \alpha^k (z^{k+1} - \nabla_w u^{k+1})$ .

The above algorithm can be expressed as the following simple form:

$$\begin{cases} u^{k+1} = u^k - \delta(A^T A u^k - A^T b + \operatorname{div}_w p^k), \\ z^{k+1} = \mathit{shrink}(\nabla_w u^{k+1} - \frac{p^k}{\alpha^k}, \frac{\mu}{\alpha^k}), \\ p^{k+1} = p^k + \alpha^k(z^{k+1} - \nabla_w u^{k+1}). \end{cases} \quad (31)$$

Hence the LPAMA does not need to solve nonlocal TV denoising subproblems and does not require to evaluate the inverses of the operator involving  $\Delta_w$  (it is required for the BOSSB and the IADM) and  $A^T A$  (it is required for the PDHG algorithm). These are advantages of our proposed algorithm LPAMA over the algorithms that we present in Subsection 2.1.

Similarly, by linearizing  $\frac{\alpha}{2}\|z^k - \nabla_w u\|_2^2$  at the first step in the framework of IADM (20), we can avoid evaluating the inverse of  $\frac{1}{\delta} - \alpha\Delta_w$  as follows:

$$\begin{cases} u^{k+1} = \arg \min_u \langle p^k, z^k - \nabla u \rangle + \langle A^T (A u^k - b), u - u^k \rangle \\ \quad + \alpha \langle \operatorname{div}_w (z^k - \nabla_w u^k), u - u^k \rangle + \frac{1}{2\delta} \|u - u^k\|_2^2 \\ z^{k+1} = \arg \min_z \mu \|z\| + \langle p^k, z - \nabla u^{k+1} \rangle + \frac{\alpha}{2} \|z - \nabla u^{k+1}\|_2^2 \\ p^{k+1} = p^k + \alpha(z^{k+1} - \nabla u^{k+1}), \end{cases} \quad (32)$$

where  $\alpha > 0$  and  $\delta > 0$ . In a more simplified form, we have

$$\begin{cases} u^{k+1} = u^k - \delta((A^T A - \alpha\Delta_w)u^k - A^T b + \operatorname{div}_w(\alpha z^k + p^k)), \\ z^{k+1} = \mathit{shrink}(\nabla u^{k+1} - \frac{p^k}{\alpha}, \frac{\mu}{\alpha}), \\ p^{k+1} = p^k + \alpha(z^{k+1} - \nabla u^{k+1}). \end{cases} \quad (33)$$

The global convergence of the above algorithm can be followed by the analysis in [36] as did for the IADM if  $0 < \delta < 1/\|A^T A - \alpha\Delta_w\|_2^2$ . We call this extended version of the IADM as ‘‘IADM-e’’.

#### 4. Algorithm Framework for Motion Deblurring and Stability

In this section, we give the details of motion deblurring process and analyze the stability of the LPAMA. We also describe how we update the weight and how to choose the parameters for algorithms. First of all, we give the following general algorithm framework for the non-blind motion deblurring problem.

In Algorithm 1, we project image  $u^m$  onto  $[0, 255]^n$  as did in [33]. By doing this, we can reduce boundary artifacts and achieve stability of the deblurring algorithms on various input images and blurring kernels.

---

**Algorithm 1** Non-blind Motion Deblurring Algorithm Framework

---

**Input:** Given blurred and noisy image  $b$  and blurring operator  $A$  with the periodic or reflexive boundary condition

**Initialization:**  $u^{-1} = 0$  and  $u^0 = \arg \min_{0 \leq v \leq 255} \|v - A^T b\|_2$

**while**  $m < K$  and  $\|u^m - u^{m-1}\|_2 > \varepsilon \|u^m\|_2$  **do**

    If  $m \bmod W = 0$ , update weight  $w(x, y) = e^{-\frac{d_{u^m}^2(x, y)}{2h^2}}$  with distance in (6).

    Update  $u^{m+1}$  by LPAMA, BOSSB, PDHG, or IADM.

    Project  $u^{m+1} = \arg \min_{0 \leq v \leq 255} \|v - u^{m+1}\|_2$ .

**end while**

---

#### 4.1. Stability Analysis for the Proposed Algorithm

In this subsection, we analyze the stability of the proposed algorithm, LPAMA. As we commented in Subsection 2.1, the BOSSB and the IADM have theoretical bounds on the parameter  $\delta$  for the convergence of them. For both the BOSSB and the IADM,  $0 < \delta < 1/\|A^T A\|_2$ . Unfortunately, the theoretical bound on the parameter  $\delta$  has not been given for the algorithms such as the LPAMA and the PDHG. Therefore, we empirically analyze the bound of the parameter  $\delta$ , which is crucial for the stability, of the LPAMA on four different images and kernels with the periodic boundary condition. Four different images and kernels are “Babara” image with  $5 \times 5$  kernel, “House” image with  $9 \times 9$  kernel, “Lena” image with  $13 \times 13$  kernel, and “Camera” image with  $17 \times 17$  kernel. We note that the PSNRs of given blurred images are 23.3dB, 23.8dB, 26.9dB, and 20.3dB respectively. The PSNR (peak signal-to-noise ratio) is defined by

$$10 \log_{10} \left( \frac{255^2 n}{\|u - \tilde{u}\|_2^2} \right)$$

where  $n$  is the size of the image,  $u$  is the original image, and  $\tilde{u}$  is the recovered image. The values of  $\|A^T A\|_2$  are  $4.3 \times 10^3$ ,  $18.7 \times 10^3$ ,  $9.9 \times 10^3$ , and  $0.8 \times 10^3$  respectively.

We terminate the algorithms when

$$\|u^k - u^{k-1}\|_2 \leq \varepsilon \|u^k\|_2 \quad (34)$$

is satisfied with the given  $\varepsilon = 10^{-4}$ . We set the maximum number of iterations as 1000

$\delta$	Babara		House		Lena		Camera	
	PSNR	Iter	PSNR	Iter	PSNR	Iter	PSNR	Iter
<b>0.00001</b>	22.7	81	23.8	57	25.5	83	19.4	87
<b>0.0001</b>	22.7	80	23.8	60	25.5	82	19.4	75
<b>0.001</b>	22.8	78	23.8	61	25.7	77	19.5	75
<b>0.01</b>	24.0	205	25.3	197	27.7	215	21.1	373
<b>0.1</b>	26.1	146	29.1	193	30.9	189	23.8	398
<b>0.5</b>	27.1	120	30.4	207	32.2	323	24.9	400
<b>1.0</b>	<i>27.4</i>	<i>124</i>	<i>30.6</i>	<i>221</i>	<i>32.5</i>	<i>338</i>	<i>25.0</i>	<i>361</i>
<b>1.5</b>	27.4	123	30.6	220	32.6	330	25.1	320
<b>1.8</b>	27.5	125	30.7	215	32.7	324	25.1	299
<b>2.0</b>	27.5	1000	30.7	1000	32.7	1000	25.0	1000
<b>2.1</b>	Unstable		Unstable		Unstable		Unstable	

Table 1: Numerical results for the various  $\delta$  (The LPAMA is stable when  $0 < \delta < 2$ ). When  $\delta > 2$ , the relative error is getting larger as the iteration increases. It is reasonable to choose  $\delta$  in the interval  $(0.1, 2)$  to get a better performance. We fix the weight ( $w = w_0$ ) for experiments.

In Table 1, we report the PSNR and the number of iterations of the LPAMA with a fixed weight for various  $\delta$  on four images. We note that the parameter  $\delta$  can be considered as a stepsize. From Table 1, the LPAMA is stable when  $0 < \delta < 2$ . However, the PSNR of the obtained solution is lower than that of the given blurred image when  $\delta < 0.01$ . If we set  $\varepsilon = 10^{-8}$  in (34) and the maximum iteration as 400,000 then we have PSNR 24.9dB at the maximum iteration when  $\delta = 0.0001$ . This shows that the algorithm converges slowly when  $\delta$  is small. When  $\delta > 2$ , the relative error is getting larger as the iteration increases. Therefore, it is reasonable to choose  $\delta$  in the interval  $(0.1, 2)$  to get a better performance.

#### 4.2. Weight Update for Nonlocal TV

To achieve better PSNR with less CPU time, we need to choose an appropriate size of the search window  $\Omega_w(x)$  and the interval of updating weights ( $W$  in Algorithm 1). In Figure 2, we report the relation between the PSNR and the size of search window and the relation between the CPU time and the size of search window when the LPAMA is applied to solve a motion deblurring problem with the ‘‘House’’ image using  $5 \times 5$  motion blurring kernel with the reflexive boundary condition. From Figure 2, we see that if the size of search window is larger than  $7 \times 7$  then the increasing rate of the PSNR becomes smaller but the CPU time is rather larger. Also, as commented in

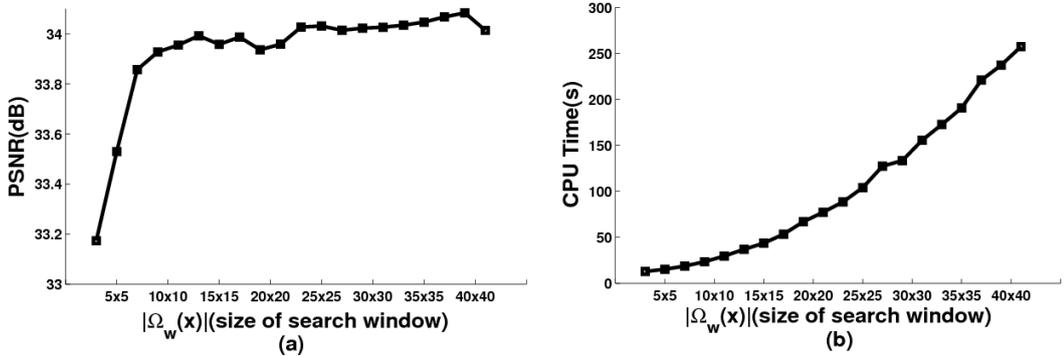


Figure 2: (a) Relation between the PSNR and the size of search window  $|\Omega_w(x)|$ , (b) Relation between the CPU time and the size of search window  $|\Omega_w(x)|$ . We note that  $K = 50$ ,  $W = 10$  in Algorithm 1, and we use the “House” image using  $5 \times 5$  motion blurring kernel with the reflexive boundary condition.

[10], when the image is not very periodic, such as the “Cameraman” image, the search window larger than  $11 \times 11$  do not show any improvement of the PSNR. Therefore, we use  $7 \times 7$  search window for our numerical experiments to reduce the CPU time with having a reasonably good PSNR. Hence we can reduce the computational burden for updating weights (for the “House” image, it is  $1.7s$  per each weight update) with a reasonably good PSNR.

In Figure 3, we present how the PSNR and the relative error are changed depending on number of iterations  $K$  for each given  $W$  when the LPAMA is applied to solve a motion deblurring problem with the “Barbara” image using  $5 \times 5$  motion blurring kernel with the reflexive boundary condition. As shown in Figure 3 (a), 2 or 3 updates of the weight dramatically increase the PSNR (for the “Barbara”, more than 2dB). Hence, in our experiments, we set  $W = 10$  for the LPAMA, the IADM, and the PDHG. Strangely, the BOSSB does not show any improvement of the PSNR with updating weights (see also [13]). Hence, we fix the weight for the BOSSB (i.e., there is no weight update) for the numerical experiments in Section 5. Since we solve a deblurring problem with a different regularizer when we update a weight, the relative error is increasing abruptly and so it does not decrease monotonically. Therefore, the relative error of the nonlocal TV deblurring problem with updating a weight is usually larger than that with the fixed weight; see Figure 3 (b).

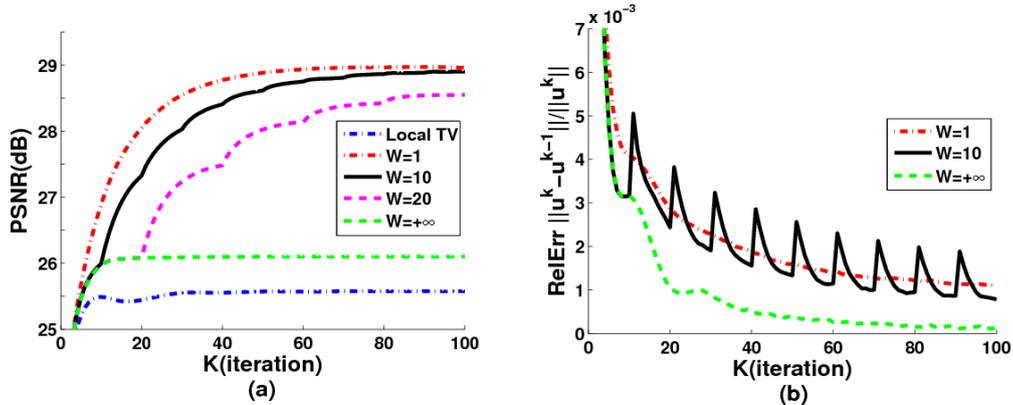


Figure 3: (a) Relation between the PSNR and the number of iterations  $K$  for various weight update rates  $W$ . (b) Relation between the relative error and the number of iterations  $K$  for various weight update rates  $W$ . We note that we use the “Barbara” image using  $5 \times 5$  motion blurring kernel with the reflexive boundary condition.

In case of local TV, if we choose a large  $\mu$ , then an edge information tends to be preserved well but the textured regions tend to be flattened. On the other hand, if we choose a small  $\mu$ , small scale features have a tendency to be relatively well preserved. Hence the local TV with a fixed regularization parameter shows poor performance. Hence, inspired by spatially varying regularization methods [8, 9], we use a heuristic of an increasing sequence of regularization parameters started from a small positive value to get better performance. In other words, if the problem (11) is to be solved with the target parameter value  $\mu = \bar{\mu}$ , we propose to solve a sequence of problems (11) defined by an increasing sequence  $\{\mu^0, \mu^1, \dots, \mu^\ell = \bar{\mu}\}$  with a given finite positive integer  $\ell$ . When a new problem, associated with  $\mu^{j+1}$  is to be solved, the approximate solution for the current problem with  $\mu = \mu^j$  is used as the starting point. In our numerical experiments, we update  $\mu^k = 10^{(k-K)/K} \mu$  at iteration  $k$ . We call this method as “LPAMA-h”. For  $\ell_1$  regularization problem, a similar methodology is known as the homotopy strategy (also known as continuation strategy [37]).

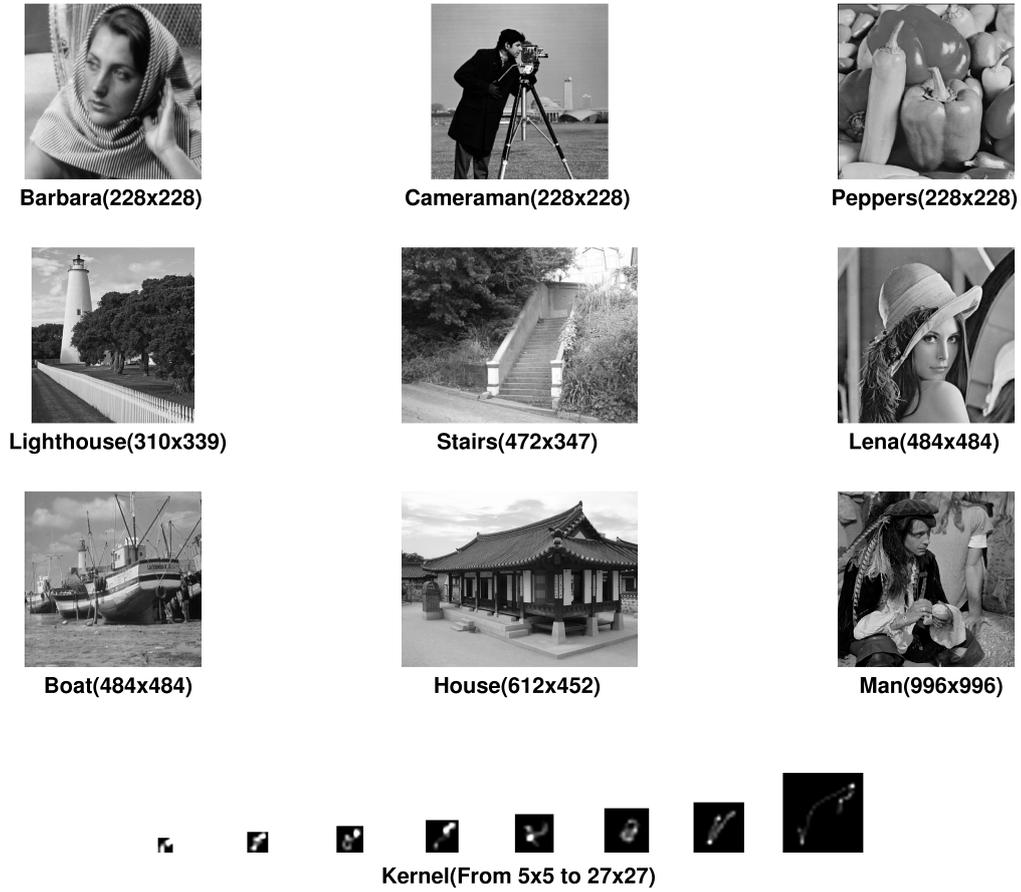


Figure 4: Dataset for the numerical experiments - 9 different test images; the sizes of image are various from  $228 \times 228$  to  $996 \times 996$ . 8 different kernels; from left to right, the size of each kernel is  $5 \times 5$ ,  $7 \times 7$ ,  $9 \times 9$ ,  $11 \times 11$ ,  $13 \times 13$ ,  $15 \times 15$ ,  $17 \times 17$  and  $27 \times 27$ .

## 5. Numerical Experience on Nonlocal TV Deblurring Problems

All algorithms are implemented with 64bit MATLAB (version 7.10). All runs are performed on a laptop with Intel i7-640LM CPU (2.13 - 2.93GHz)

$\delta$	B.C.	REG.	BOSSB	PDHG	LPAMA	LPAMA-h	IADM	IADM-e
			PSNR/TIME	PSNR/TIME	PSNR/TIME	PSNR/TIME	PSNR/TIME	PSNR/TIME
0.5	Reflexive	NLTV	28.9/54.4	28.3/114.1	28.8/ <b>23.9</b>	<b>29.0</b> /24.0	28.7/27.5	28.7/24.3
		LTV	<b>28.5</b> /13.4	27.5/52.9	27.8/ <b>6.9</b>	28.3/ <b>6.9</b>	27.3/7.2	27.3/ <b>6.9</b>
	Periodic (FFT)	NLTV	27.4/63.9	26.9/43.9	27.4/ <b>39.8</b>	<b>27.6</b> /40.3	27.4/43.2	27.4/40.4
		LTV	<b>27.3</b> /23.4	26.4/18.9	26.7/16.5	27.1/16.5	26.3/16.9	26.3/ <b>16.2</b>
1.0	Reflexive	NLTV	28.8/54.2	29.4/115.2	29.7/ <b>24.0</b>	<b>29.9</b> / <b>24.0</b>	29.6/27.3	29.6/24.3
		LTV	28.5/13.5	28.3/53.9	28.5/6.9	<b>29.1</b> /6.9	27.8/7.1	27.8/ <b>6.8</b>
	Periodic (FFT)	NLTV	26.8/64.4	28.0/43.9	28.4/ <b>39.8</b>	<b>28.5</b> / <b>39.8</b>	28.3/43.0	28.3/40.0
		LTV	26.9/23.1	27.2/19.1	27.5/ <b>16.4</b>	<b>27.9</b> /16.5	26.9/16.8	27.0/ <b>16.4</b>
1.5	Reflexive	NLTV	10.3/54.4	29.7/114.6	29.7/ <b>24.1</b>	<b>29.9</b> / <b>24.1</b>	29.7/27.5	29.7/24.5
		LTV	10.2/13.6	28.6/53.4	28.6/ <b>6.8</b>	<b>29.2</b> /6.9	28.0/7.1	28.0/ <b>6.8</b>
	Periodic (FFT)	NLTV	10.3/64.3	28.4/44.4	28.6/39.8	<b>28.8</b> / <b>39.5</b>	28.6/43.3	28.6/40.3
		LTV	10.1/23.4	27.5/19.1	27.7/16.5	<b>28.2</b> / <b>16.4</b>	27.2/16.8	27.2/16.5

Table 2: Comparison of the performance of the LPAMA(-h) with that of the BOSSB, the PDHG, and the IADM(-e) for three different choices of  $\delta$ . The BOSSB is stable and has better performance results when  $\delta = 0.5, 1.0$ . On the other hand, the PDHG, the LPAMA(-h), and the IADM(-e) are stable and have better performance results when  $\delta = 1.0, 1.5$ . It is reasonable to choose  $\delta = 1$  for all the algorithms. NLTV stands for nonlocal TV and LTV stands for local TV. The LPAMA-h overall outperforms other algorithms.

and 8GB Memory. The Operation System is 64bit Linux. We note that we slightly modified the BOSSB (<http://www.math.ucla.edu/~xqzhang/html/code.html>) to enhance the performance for the motion deblurring problems, and the IADM and the PDHG algorithms are implemented by us based on the guideline in [19] and [20] respectively. In order to speed up the computation of weights and nonlocal operators, such as  $\text{div}_w$  and  $\nabla_w$ , we compute them in C language with interface to MATLAB through the mex function. All the test images are blurred and noised. The test images and blurring kernels are in Figure 4. We use 9 different images. The size of tested images is various from  $228 \times 228$  to  $996 \times 996$ . Also, we use 8 different motion blurring kernels. Some of them are freely available from <http://www.wisdom.weizmann.ac.il/~levina/>. The size of kernels varies from  $5 \times 5$  to  $27 \times 27$ ; see Figure 4 for details. We fix the added Gaussian random noise,  $\eta \sim \mathcal{N}(0, 9)$  in all experiments. Also we set  $5 \times 5$  patch  $B(x)$  and  $h = 2\sigma$  where  $\sigma$  is the estimated noise level (we set  $\sigma = 3$ ). Hence we have 72 different blurred and noisy images for our experiments.

Table 2 reports the mean values of the PSNR and the CPU time of 72 instances for three different choices of  $\delta$ . Since the boundary artifact is much less in the reflexive boundary condition, the performance of all algorithms are

better when the reflexive boundary condition is used. The BOSSB is more sensitive to the boundary condition. For nonlocal TV, the LPAMA-h obtains best PSNR and the LPAMA is faster than other algorithms. For local TV with  $\delta = 0.5$ , the BOSSB obtains better PSNR than other algorithms. For local TV with  $\delta = 1.0$  or  $1.5$ , the LPAMA-h obtains better PSNR than other algorithms. The LPAMA(-h) is comparable with the IADM-e in terms of the CPU time for local TV. The BOSSB is stable and has better performance results when  $\delta = 0.5, 1.0$ . On the other hand, the PDHG, the LPAMA, and the IADM are stable and have better performance results when  $\delta = 1.0, 1.5$ . It is reasonable to choose  $\delta = 1$  for all the algorithms.

As we mentioned in Section 4, the heuristic of using an increasing sequence of regularization parameters helps to enhance the image quality, especially for local TV. But, since the nonlocal TV updates the weight regularly, this heuristic does not have advantage for nonlocal TV. We see that the PSNR of the LPAMA-h is better than that of the LPAMA especially for local TV.

In the following, we describe how we set the parameters for all the algorithms. For the BOSSB [13], we set 30 outer iterations (i.e.,  $K = 30$ ) and 10 inner iterations with  $\delta = 1, \mu = 50, \alpha = 200$ . Since the parameters used in [13] does not work well for our motion deblurring testing problems, we modified the parameters for the better performance. For the PDHG [20], we set 50 iterations (i.e.,  $K = 50$ ) with  $\delta = 1, \mu = 1.1, \tau = 0.495$ . For the reflexive boundary condition, we use conjugate gradient method (up to 5 iterations) to evaluate the inverse of the operator  $\frac{1}{\delta}I + A^T A$ . For the periodic boundary condition, we use FFT to compute the inverse of the operator  $\frac{1}{\delta}I + A^T A$ . Adaptive parameter update scheme has been proposed in [20], but this scheme does not work well on our motion deblurring problems. Hence we fixed parameters for the PDHG algorithm as did in [28]. We further tuned the parameters for the better performance. For the LPAMA(-h) and the IADM(-e), we set 50 iterations (i.e.,  $K = 50$ ) with  $\delta = 1, \mu = 1.1, \alpha = 50$ . All the parameters for the LPAMA(-h) and the IADM(-e) were tuned for the better performance. For local TV deblurring problems, we use the same parameters as those of algorithms for nonlocal TV except for  $K$ . Since we do not need to update the weight for local TV, we set  $K = 30$  for all the algorithms; see Figure 3 (a).

In Table 3, we report the PSNR and the CPU time of all the algorithms for solving local TV and nonlocal TV deblurring problems on 9 different images. The PSNR and the CPU time in Table 3 are the mean value of them on 8 different motion blurring kernels for each algorithm. When the images,

B.C.	REG.	IMAGE	BOSSB	PDHG	LPAMA	LPAMA-h	IADM	IADM-e
			PSNR/TIME	PSNR/TIME	PSNR/TIME	PSNR/TIME	PSNR/TIME	PSNR/TIME
Reflexive	NLTV	Barbara	25.9/11.6	27.3/26.5	27.9/ <b>4.8</b>	<b>28.5/4.8</b>	27.7/5.5	27.7/4.9
		Cameraman	28.3/11.5	28.8/25.5	<b>29.0/5.0</b>	<b>29.0/5.0</b>	28.9/5.8	28.9/5.1
		Pepper	29.7/11.6	<b>30.3/26.2</b>	30.2/ <b>5.0</b>	30.1/5.1	<b>30.3/5.7</b>	<b>30.3/5.0</b>
		LightHouse	26.2/30.2	26.7/48.9	27.1/11.1	<b>27.4/10.9</b>	27.0/13.0	27.0/11.2
		Stairs	26.9/43.1	27.3/77.4	27.7/16.6	<b>28.1/16.4</b>	27.6/19.7	27.6/17.3
		Lena	32.5/52.9	32.9/105.1	<b>33.1/22.8</b>	<b>33.1/22.6</b>	<b>33.1/26.0</b>	<b>33.1/23.1</b>
		Boat	29.8/52.4	30.3/105.7	<b>30.6/22.4</b>	<b>30.6/21.8</b>	30.5/25.8	30.5/22.8
		House	29.6/63.0	30.1/125.7	30.5/ <b>26.2</b>	<b>30.7/26.6</b>	30.4/31.1	30.4/27.4
		Man	30.9/211.3	31.1/496.2	31.4/102.4	<b>31.6/103.4</b>	31.4/112.8	31.4/ <b>102.3</b>
		<b>AVERAGE</b>	<b>28.8/54.2</b>	<b>29.4/115.2</b>	<b>29.7/24.0</b>	<b>29.9/24.0</b>	<b>29.6/27.3</b>	<b>29.6/24.3</b>
	LTV	Barbara	25.3/2.4	25.6/12.3	25.8/ <b>1.5</b>	<b>26.9/1.5</b>	25.3/ <b>1.5</b>	25.3/ <b>1.5</b>
		Cameraman	27.9/2.4	26.9/12.3	27.1/ <b>1.4</b>	<b>27.8/1.5</b>	25.8/ <b>1.4</b>	25.8/ <b>1.4</b>
		Pepper	29.8/2.4	29.8/12.5	29.9/ <b>1.5</b>	<b>30.0/1.5</b>	29.1/ <b>1.5</b>	29.2/ <b>1.5</b>
		LightHouse	25.8/4.9	25.6/22.2	25.8/ <b>3.0</b>	<b>26.4/3.0</b>	25.0/3.1	25.1/ <b>3.0</b>
		Stairs	26.9/8.1	26.8/34.2	27.0/4.6	<b>27.7/4.6</b>	26.6/4.6	26.6/ <b>4.3</b>
		Lena	32.1/12.5	32.1/46.9	32.2/6.0	<b>32.5/6.0</b>	31.8/6.6	31.8/ <b>5.9</b>
		Boat	29.5/12.0	29.1/49.1	29.3/6.3	<b>29.8/5.9</b>	28.6/6.2	28.6/6.2
		House	28.9/14.9	28.6/59.7	28.7/7.7	<b>29.3/7.7</b>	28.0/7.9	28.0/ <b>7.5</b>
		Man	30.7/62.2	30.4/235.9	30.6/ <b>30.1</b>	<b>31.0/30.6</b>	30.0/31.3	30.1/30.2
		<b>AVERAGE</b>	<b>28.5/13.5</b>	<b>28.3/53.9</b>	<b>28.5/6.9</b>	<b>29.1/6.9</b>	<b>27.8/7.1</b>	<b>27.8/6.8</b>
Periodic (FFT)	NLTV	Barbara	24.7/12.2	25.8/7.1	26.4/ <b>6.4</b>	<b>26.7/6.4</b>	26.3/7.3	26.3/6.5
		Cameraman	26.2/12.1	27.4/7.1	<b>27.7/6.4</b>	<b>27.7/6.4</b>	<b>27.7/7.1</b>	<b>27.7/6.5</b>
		Pepper	26.1/12.2	27.1/7.1	<b>27.4/6.4</b>	<b>27.4/6.5</b>	<b>27.4/7.2</b>	<b>27.4/6.6</b>
		LightHouse	24.3/33.2	25.5/18.0	25.8/16.3	<b>26.1/16.1</b>	25.7/18.2	25.7/16.6
		Stairs	24.2/52.5	25.8/31.8	26.2/28.6	<b>26.5/28.3</b>	26.1/31.7	26.1/29.1
		Lena	29.5/54.2	31.4/25.3	31.7/23.9	<b>31.8/23.8</b>	31.7/27.4	31.7/24.1
		Boat	28.7/54.2	29.5/25.6	<b>29.8/23.9</b>	<b>29.8/23.9</b>	29.7/27.7	29.7/24.5
		House	27.9/69.4	29.3/41.5	29.7/ <b>37.0</b>	<b>29.9/37.4</b>	29.6/41.4	29.6/38.1
		Man	29.8/279.4	30.5/231.8	30.9/209.5	<b>31.1/209.3</b>	30.8/218.6	30.9/ <b>208.0</b>
		<b>AVERAGE</b>	<b>26.8/64.4</b>	<b>28.0/43.9</b>	<b>28.4/39.8</b>	<b>28.5/39.8</b>	<b>28.3/43.0</b>	<b>28.3/40.0</b>
	LTV	Barbara	24.4/3.3	24.8/2.7	25.0/ <b>2.2</b>	<b>25.7/2.2</b>	24.6/2.3	24.6/ <b>2.2</b>
		Cameraman	26.3/3.2	26.0/2.7	26.3/2.3	<b>26.8/2.3</b>	25.2/ <b>2.2</b>	25.2/ <b>2.2</b>
		Pepper	26.5/3.3	26.8/2.7	27.2/ <b>2.2</b>	<b>27.3/2.3</b>	26.8/2.3	26.9/2.3
		LightHouse	24.4/8.1	24.6/7.1	24.9/6.1	<b>25.4/6.0</b>	24.3/6.2	24.3/ <b>6.0</b>
		Stairs	24.6/15.3	25.5/14.0	25.7/12.3	<b>26.2/12.0</b>	25.4/12.4	25.5/12.1
		Lena	29.7/13.2	30.8/8.1	31.1/ <b>7.2</b>	<b>31.3/7.3</b>	30.8/7.5	30.8/ <b>7.2</b>
		Boat	28.6/13.4	28.4/8.2	28.7/ <b>7.1</b>	<b>29.1/7.2</b>	28.1/7.5	28.1/7.3
		House	27.6/21.6	28.0/16.3	28.2/ <b>14.1</b>	<b>28.7/14.2</b>	27.6/14.5	27.6/14.2
		Man	29.8/126.6	29.9/110.4	30.1/ <b>94.0</b>	<b>30.5/95.4</b>	29.7/96.3	29.7/94.5
		<b>AVERAGE.</b>	<b>26.9/23.1</b>	<b>27.2/19.1</b>	<b>27.5/16.4</b>	<b>27.9/16.5</b>	<b>26.9/16.8</b>	<b>27.0/16.4</b>

Table 3: Comparison of the performance of the LPAMA(-h) with that of BOSSB, PDHG and IADM(-e) for each different image. For each image, the PSNR (dB) and the CPU Time (sec) are the average values of the deblurred results on eight different blurring kernels.

such as the “Barbara” and the “House”, have the strong periodic structure, the PSNR of the recovered images by using nonlocal TV is much better than that by using local TV; see also Figures 5 and 6. From Table 3, we see that the LPAMA-h obtains better PSNR than all the other algorithms on average. The LPAMA(-h) is at least two times faster than both the PDHG and the BOSSB and slightly faster than the IADM when the reflexive boundary condition is used. The LPAMA(-h) is faster than all the other algorithms for nonlocal TV when FFT is used with the periodic boundary condition. The BOSSB gets the PSNR that is at least 1.1dB lower than that of the LPAMA-h for nonlocal TV. We think that, for the BOSSB, the reason of obtaining much less PSNR is caused by fixing the weight; see also Figure 3 (a).

In Figures 5-9, the deblurred images by using the LPAMA-h and other algorithms are shown along with the PSNR and the CPU time. We note that the reflexive boundary condition is used except for the “Cameraman” image.

Figure 5 shows the results of  $228 \times 228$  “Barbara” image with  $15 \times 15$  kernel. When the deblurred images by using nonlocal TV is compared with those by using local TV, it is shown that the weak textures are recovered with little artifacts, especially when the LPAMA-h is applied. It is surprising that the LPAMA-h with local TV obtains a better PSNR than the BOSSB with nonlocal TV. The reason is clear since the BOSSB does not update weight for nonlocal TV. Hence it does not take advantage of using nonlocal TV. On the other hand, LPAMA-h adaptively changes a regularization parameter  $\mu$  to recover better textures as explained at the end of Section 4.

Figure 6 shows the results of  $612 \times 452$  “House” image with  $27 \times 27$  kernel. The LPAMA-h with nonlocal TV is more than two times faster than the PDHG with local TV and also has a much better PSNR. The LPAMA-h is faster than the IADM for nonlocal TV. The IADM is comparable with the LPAMA-h in terms of the CPU time for local TV, but IADM gets a lower PSNR than the PDHG and BOSSB. Since the weight update scheme does not work in BOSSB, the advantage of nonlocal TV is small even though the image has periodic patterns. Hence the PSNR of the LPAMA-h with local TV is comparable to that of the BOSSB with nonlocal TV.

Figure 7 shows the results of  $484 \times 484$  “Lena” image with  $7 \times 7$  kernel. It is shown that the LPAMA-h is at least four times faster than other algorithms except the IADM when it is applied to local TV. But the IADM obtains a lower PSNR for local TV. We can see that the hair part in the upper right on the deblurred “Lena” image by the LPAMA-h (see Figure 7 (d)) is very



Figure 5: Comparison on the deblurred “Barbara” image ( $228 \times 228$ ) using  $15 \times 15$  kernel with the reflexive boundary condition. When the deblurred images by using nonlocal TV is compared with those by using local TV, it is shown that the weak textures are recovered with little artifacts, especially when the LPAMA-h is applied. Since the BOSSB does not update weight for nonlocal TV, it does not take advantage of using nonlocal TV.

well recovered when it compared with other deblurred images.

Figure 8 shows the results of  $484 \times 484$  “Boat” image with  $5 \times 5$  kernel.

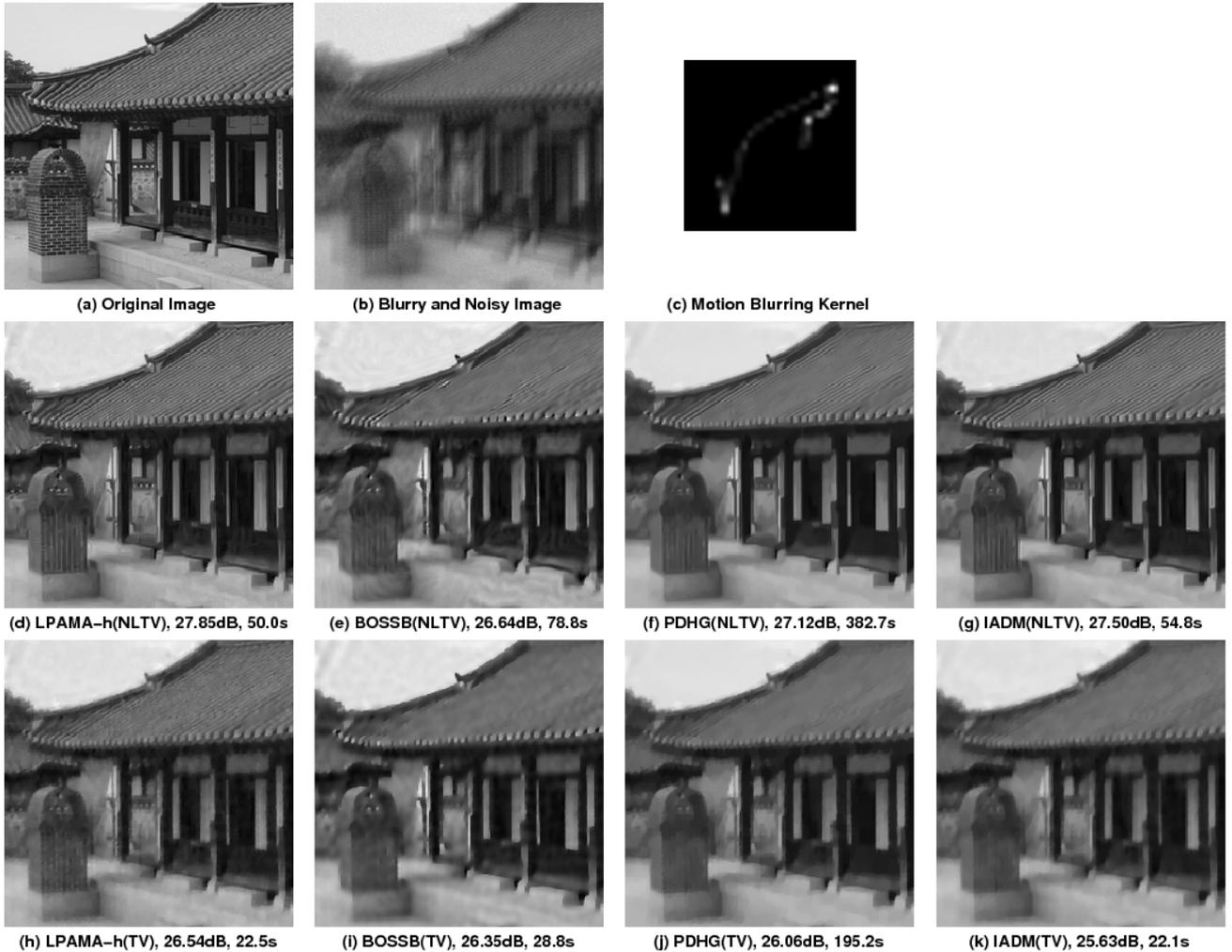


Figure 6: Comparison on the deblurred “House” image ( $612 \times 452$ ) using  $27 \times 27$  kernel with the reflexive boundary condition. We note that images are cropped. The LPAMA-h adaptively changes a regularization parameter  $\mu$  and so the deblurred image by it has better textures, especially for local TV. Hence the PSNR of the LPAMA-h is comparable to that of the BOSSB with nonlocal TV.

It is shown that the LPAMA-h is the faster than other algorithms. Even though, for nonlocal TV, the PSNR of the LPAMA-h is comparable to that



Figure 7: Comparison on the deblurred “Lena” image ( $484 \times 484$ ) using  $7 \times 7$  kernel with the reflexive boundary condition. We note that images are cropped. The hair part in the upper right on the deblurred “Lena” image by the LPAMA-h (see (d)) is very well recovered when it compared with other deblurred images.

of the IADM. The parallel wires in the lower right part of the recovered image of the LPAMA-h can be more distinguishable than that of other algorithms.

Figure 9 shows the results of  $228 \times 228$  “Cameraman” image with  $11 \times 11$

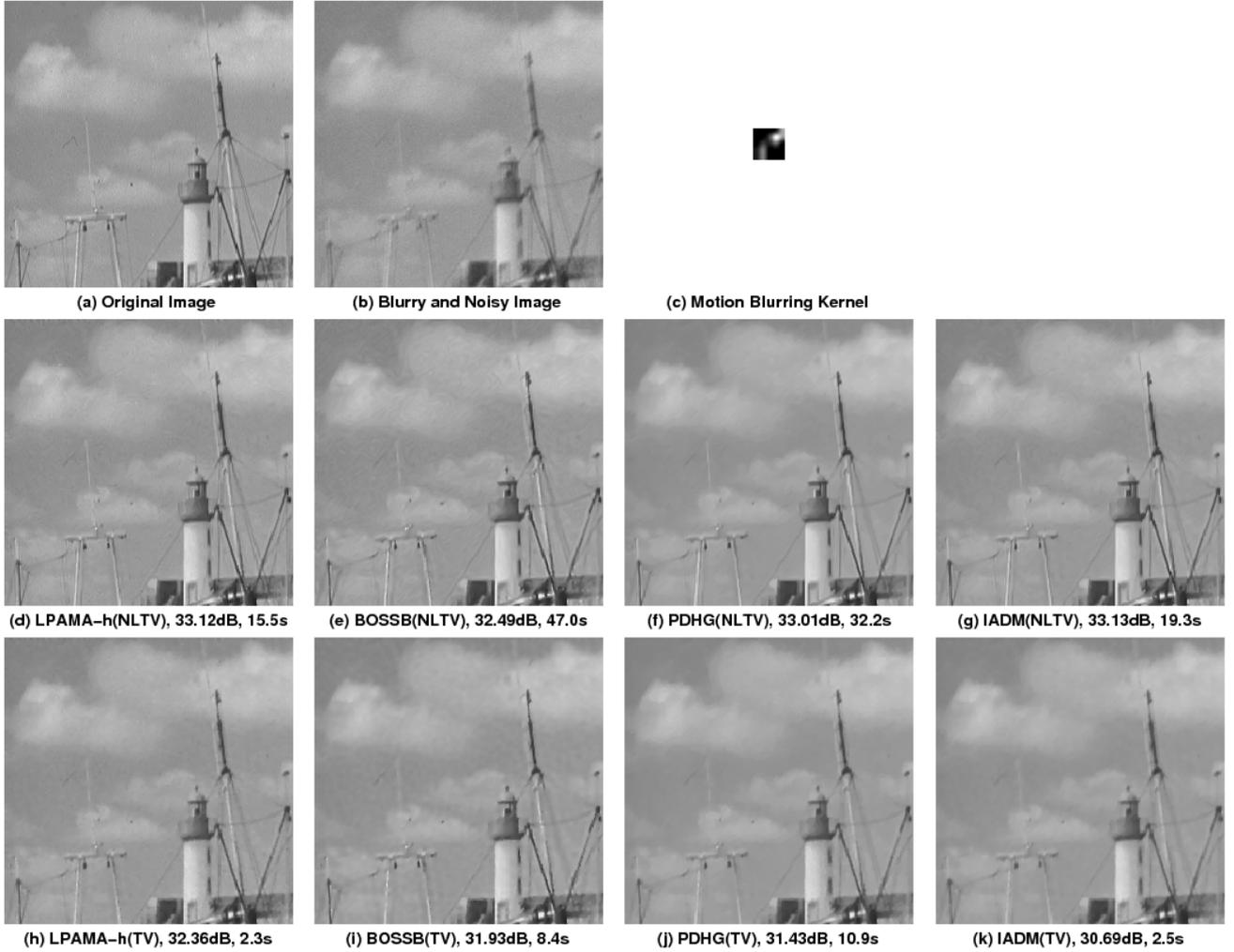


Figure 8: Comparison on the deblurred “Boat” image ( $484 \times 484$ ) using  $5 \times 5$  kernel with the reflexive boundary condition. We note that images are cropped. The parallel wires in the lower right part of the recovered image of the LPAMA-h can be more distinguishable than that of other algorithms.

kernel. For this image, FFT is used with the periodic boundary condition. Even though the kernel size is relatively small, the restored image by the BOSSB shows strong boundary artifacts. We see that the LPAMA-h is faster than other algorithms except the IADM when it is applied to local TV. It

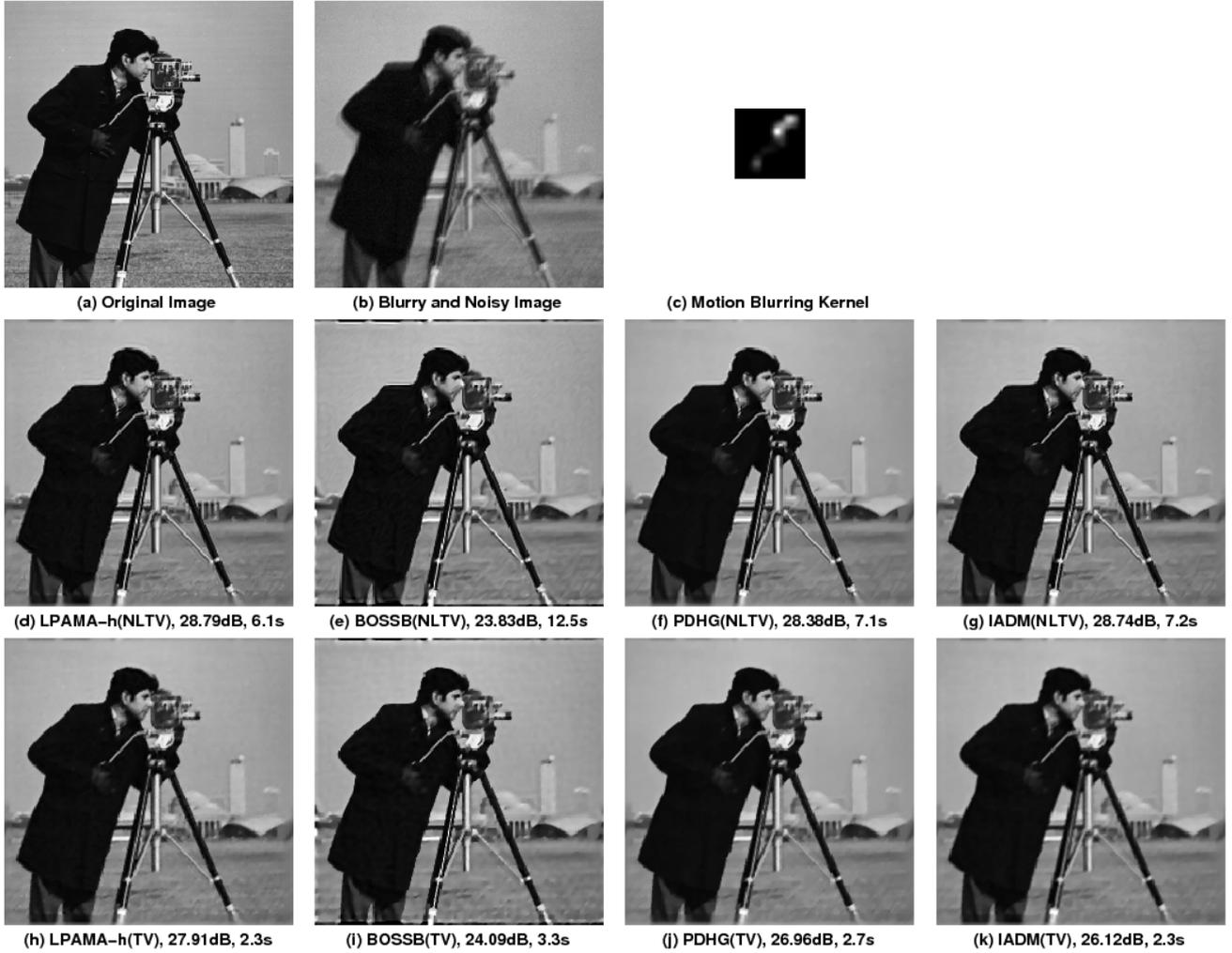


Figure 9: Comparison on the deblurred “Cameramen” image ( $228 \times 228$ ) using  $11 \times 11$  kernel. FFT is used with the periodic boundary condition. Even though the kernel size is relatively small, the restored image by the BOSSB shows strong boundary artifacts.

achieves better PSNRs than other algorithms for both nonlocal and local TV.

## 6. Conclusion

In this paper we have proposed the linearized proximal alternating minimization algorithm for solving motion deblurring problems based on nonlocal total variation. The linearized proximal alternating minimization algorithm has advantages of avoiding inner loops and the computation of any inverses involving the blurring operator and the nonlocal operator by using the linearization of the fidelity term and the proximal function. We have compared our method with the Bregmanized operator splitting with the splitting Bregman method [13], the inexact alternating direction method [19], and the primal dual hybrid gradient algorithm [20]. The numerical simulation results show that our algorithm overall outperforms all the other algorithms and so is efficient and robust for nonlocal total variation based motion deblurring problems. From our numerical experiments, we see that if the weight update scheme does not work, then the advantage of nonlocal TV is very small even though the image has lots of periodic patterns.

Our proposed algorithm can be extended to solve more general separable convex optimization problem (22), whose objective function is the sum of a smooth function and a nonsmooth function. Hence this algorithm can be viewed as a generalization of the alternating minimization algorithm proposed by Tseng [21] and the relaxed AMA described in [28]. The theoretical analysis of the convergence properties of the linearized proximal alternating minimization algorithm for solving separable convex minimization problems has not been done yet. This can be a future research topic.

## Acknowledgments

We would like to thank the anonymous reviewers for their valuable suggestions for improving this paper. This work was supported by Priority Research Centers Program(2009-0093827) and Basic Science Research Program(2010-0510-1-3) through the National Research Foundation of Korea funded by the Ministry of Education, Science and Technology.

## References

- [1] R. FERGUS, B. SINGH, A. HERTZMANN, S.T. ROWEIS, AND W.T. FREEMAN, Removing camera shake from a single photograph, ACM Trans. on Graphics, 25 (2006), 787-794.

- [2] Q. SHAN, J. JIA, AND A. AGARWALA, High-quality motion deblurring from a single image, *ACM Trans. on Graphics*, 27 (2008), 73:1-73:10.
- [3] C. WANG, L. SUN, Z.Y. CHEN, J.W. ZHANG, AND S.Q. YANG, Multi-scale blind motion deblurring using local minimum, 2009, to appear in *Inverse Problems*.
- [4] T.F. CHAN, AND C.K. WONG, Total variation blind deconvolution, *IEEE Trans. on Image Processing*, 7 (1998), 370-375.
- [5] A. LEVIN, Y. WEISS, F. DURAND, AND W.T. FREEMAN, Understanding and evaluating blind deconvolution algorithms, *IEEE Conf. on Computer Vision and Pattern Recognition*, 2009.
- [6] L. RUDIN, S. OSHER, AND E. FATEMI, Nonlinear total variation based noise removal algorithms. *Physica D*, 60 (1992), 259-268.
- [7] J. P. OLIVEIRA, J. M. BIOUCAS-DIAS, M A.T. FIGUEIREDO, Adaptive total variation image deblurring: A majorization-minimization approach, *Signal Processing*, 89 (2009), 1683-1693.
- [8] G. GILBOA, N. SOCHEN, AND Y.Y. ZEEVI, Variational denoising of partly textured images by spatially varying constraints, *IEEE Trans. on Image Processing*, 15 (2006), 2281-2289.
- [9] F. LI, C. SHEN, C. SHEN AND G. ZHANG, Variational denoising of partly textured images, *J. Vis. Commun. Image.*, 20 (2009), 293-300.
- [10] G. GILBOA, AND S. OSHER, Nonlocal linear image regularization and supervised segmentation, *Multiscale Model. Simul.*, 6 (2007), 595-630.
- [11] S. KINDERMANN, S. OSHER, AND P. W. JONES, Deblurring and denoising of images by nonlocal functionals, *Multiscale Model. Simul.* 4 (2005), 1091-1115.
- [12] Y. LOU, X. ZHANG, S. OSHER, A. BERTOZZI, Image recovery via nonlocal operators, *J. Sci. Comput.*, 42 (2010), 185-197.
- [13] X. ZHANG, M. BURGER, X. BRESSON AND S. OSHER, Bregman-ized nonlocal regularization for deconvolution and sparse reconstruction, *SIAM J. Imaging Sciences*, 3 (2010), 253-276.

- [14] G. PEYRE, S. BOUGLEUX, AND L.D. COHEN, Nonlocal regularization of inverse problems, Proc. of ECCV'08, D. A. Forsyth, P. H. S. Torr, and A. Zisserman (Eds.), Lecture Notes in Computer Science, Springer, 2008, 57-68.
- [15] X. BRESSON, AND T.F. CHAN, Nonlocal unsupervised variational image segmentation models, UCLA Cam-Report, 2008.
- [16] A. ELMOATAZ, O. LEZORAY, AND S. BOUGLEUX, Nonlocal discrete regularization on weighted graphs: a framework for image and manifold processing. IEEE Trans. on Image Processing, 17 (2008), 1047-1060.
- [17] G. GILBOA, AND S. OSHER, Nonlocal operators with applications to image processing, Multiscale Model. Simul., 7 (2008), 1005-1028.
- [18] Y. WANG, J. YANG, W. YIN, AND Y. ZHANG A new alternating minimization algorithm for total variation image reconstruction, SIAM J. Imaging Sci., 1 (2008), 248-272.
- [19] Y. XIAO AND J. YANG A fast algorithm for total variation image reconstruction from random projections, Preprint, 2010.
- [20] M. ZHU AND T. CHAN, An efficient primal-dual hybrid gradient algorithm for total variation image restoration, UCLA CAM-Report, 2008.
- [21] P. TSENG, Applications of a splitting algorithm to decomposition in convex programming and variational inequalities. SIAM J. Control and Optimization, 29 (1991), 119-138.
- [22] R. LIU, AND J. JIA, Reducing boundary condition artifacts in image deconvolution, Proc. of the International Conference on Image Processing, 2008, 505-508.
- [23] T.F. CHAN, AND A.M. YIP, AND F.E. PARK, Simultaneous total variation image inpainting and blind deconvolution, 15 (2005), 92-102.
- [24] M. DONATELLI AND S. CAPIZZANO, Anti-reflective boundary conditions and re-blurring, Inverse Problems, 22 (2005), 169-182.
- [25] C. HANSEN, J.G. NAGY, AND D.P. O'LEARY, Deblurring images - matrices, spectra, and filtering, SIAM, Philadelphia, 2006.

- [26] P. ARIAS, V. CASELLES, AND G. SAPIRO, A variational framework for non-local image inpainting, Proc. of the 7th International Conference on EMMCVPR, D. Cremers, Y. Boykov, A Blake, and F. R. Schmidt (Eds.), Lecture Notes in Computer Science, Springer, 2009, 345-358.
- [27] T. GOLDSTEIN, AND S. OSHER, The split Bregman method for l1 regularized problems, SIAM J. Imaging Sci., 2 (2009), 323-343.
- [28] E. ESSER, X. ZHANG, AND T. CHAN, A general framework for a class of first order primal-dual algorithms for TV minimization, UCLA CAM-Report, 2009.
- [29] X. ZHANG, M. BURGER, AND S. OSHER, A unified primal-dual algorithm framework based on Bregman iteration, UCLA CAM-Report, 2009.
- [30] M. TAO AND J. YANG, Alternating direction algorithm for total variation deconvolution in image reconstruction, Preprint, 2009.
- [31] M.V. AFONSO, J.M. BIOUCAS-DIAS, AND M.A.T. FIGUEIREDO, Fast image recovery using variable splitting and constrained optimization, Preprint, 2009.
- [32] A. CHAMBOLLE, An algorithm for total variation minimization and applications, J. Math. Imaging and Vision, 20 (2004), 89-97.
- [33] A. BECK AND M. TEBoulLE, Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems, IEEE Trans. on Image Processing, 18 (2009), 2419-2434.
- [34] A. CHAMBOLLE, V. CASELLES, M. NOVAGA, D. CREMERS, AND T. POCK, An introduction to total variation for image analysis, Preprint, 2009.
- [35] P. COMBETTES AND W. WAJS, Signal recovery by proximal forward-backward splitting, Multiscale Model. Simul., 4 (2005), 1168-1200.
- [36] B. HE, L. Z. LIAO, D. HAN AND H. YANG, A new inexact alteratin directions method for monotone variational inequalities, Math. Program., 92 (2002), 103-118.

- [37] E. T. HALE, W. YIN, AND Y. ZHANG, Fixed-point continuation for  $l_1$ -minimization: Methodology and convergence. *SIAM J. Optim.*, 19 (2008), 1107-1130.