Noname manuscript No.
(will be inserted by the editor)

Total variation image deblurring with space-varying kernel

Daniel O'Connor · **Lieven Vandenberghe**

Abstract Image deblurring techniques based on convex optimization formulations, such as total-variation deblurring, often use specialized first-order methods for large-scale nondifferentiable optimization. A key property exploited in these methods is spatial invariance of the blurring operator, which makes it possible to use the fast Fourier transform (FFT) when solving linear equations involving the operator. In this paper we extend this approach to two popular models for space-varying blurring operators, the Nagy-O'Leary model and the Efficient Filter Flow model. We show how splitting methods derived from the Douglas-Rachford algorithm can be implemented with a low complexity per iteration, dominated by a small number of FFTs.

Keywords image deblurring, total variation, convex optimization, monotone operators, Douglas-Rachford algorithm

Mathematics Subject Classification (2000) 65K05, 90C25, 49M27, 68U10

1 Introduction

In many popular approaches to non-blind image deblurring, a deblurred image is computed by solving an optimization problem of the form

$$\underset{x}{\text{minimize}} \quad \phi_{\mathrm{f}}(Kx-b)+\phi_{\mathrm{r}}(Dx)+\phi_{\mathrm{c}}(x) \tag{1}$$

where $\phi_{\mathrm{f}}, \phi_{\mathrm{r}}$, and ϕ_{c} are convex penalty or indicator functions [21,40]. Here b is a vector containing the pixel intensities of an $M \times N$ blurry, noisy image, stored as a vector of length $n = MN$, for example, in column-major order. The optimization variable $x \in \mathbb{R}^n$ is the deblurred image. The matrix K models the blurring process, and is assumed to be known. The first term in (1) is often called the "data fidelity" term and encourages $Kx \approx b$. Typical choices for the penalty function ϕ_{f} in the data fidelity term include the squared L_2 norm, the L_1 norm, and the Huber penalty. The second term in the cost function is a regularization term. In total-variation (TV) deblurring [38,39], the matrix D represents a discrete gradient operator, and ϕ_{r} is an L_1 norm or the isotropic norm

$$\left\| \begin{bmatrix} u \\ v \end{bmatrix} \right\|_{\mathrm{iso}} = \sum_{i=1}^{n} \sqrt{u_i^2 + v_i^2}. \tag{2}$$

In tight frame regularized deblurring [30,27], D represents the analysis operator for a tight frame and ϕ_{r} is the L_1 norm. The last term $\phi_{\mathrm{c}}(x)$ can be added to enforce constraints on x such as $0 \leq x \leq 1$ by choosing ϕ_{c} to be the indicator function of the constraint set.

When devising efficient methods for solving (1), a key requirement is to exploit structure in K. Most work on image deblurring has modeled the blur as being spatially invariant, in which case K represents a convolution with a spatially invariant point spread function. This allows efficient multiplication by K and K^T by means of the fast

Daniel O'Connor
Department of Mathematics, University of California Los Angeles
E-mail: daniel.v.oconnor@gmail.com

Lieven Vandenberghe
Electrical Engineering Department, University of California Los Angeles
E-mail: vandenbe@ucla.edu

Fourier transform. Additionally, linear systems involving K can be solved efficiently because K can be expressed as a product of the 2-dimensional DFT matrix, a diagonal matrix, and the inverse of the 2-dimensional DFT matrix. In combination with fast Fourier transform techniques, methods such as the Chambolle-Pock algorithm [9,44], the Douglas-Rachford algorithm [29,12,36,33], and the alternating direction method of multipliers (ADMM) or split Bregman method [19,10,1] are very effective for solving (1). By exploiting the space-invariant convolution structure, these methods achieve a very low per-iteration complexity, equal to the cost of a small number of FFTs.

In this paper we address the problem of solving (1) under the assumption of *spatially variant* blur. We consider two models of spatially variant blur: the classical model of Nagy and O'Leary [32] and the related "Efficient Filter Flow" (EFF) model [24]. The Nagy-O'Leary model expresses a spatially variant blur operator as a sum

$$K = \sum_{p=1}^{P} U_p K_p, \tag{3}$$

where each K_p represents a spatially invariant convolution operator, and the matrices U_p are nonnegative diagonal matrices that add up to the identity. The EFF model uses a model of the form

$$K = \sum_{p=1}^{P} K_p U_p, \tag{4}$$

with similar assumptions on K_p and U_p. Unlike the space-invariant convolution operator, the space-varying models (3) and (4) cannot be diagonalized using the DFT. This renders standard Douglas-Rachford-based methods for space-invariant deblurring unusable. The contribution of the paper is to present Douglas-Rachford splitting methods for solving (1), for the two spatially variant blur models, with an efficiency that is comparable with FFT-based methods for spatially invariant blurring operators. Specifically, for many common types of fidelity and regularization terms, the cost per iteration is $O(PN^2 \log N)$ for N by N images.

Our approach is to use the structure of the blur operator to write deblurring problems in the canonical form

$$\underset{x}{\text{minimize}} \quad f(x) + g(Ax) \tag{5}$$

in such a way that the convex functions f and g have inexpensive proximal operators, and linear systems involving A can be solved efficiently. Problem (5) can then be solved by methods based on the Douglas-Rachford splitting algorithm, yielding algorithms whose costs are dominated by the calculation of order P fast Fourier transforms at each iteration. The key point is to recognize the correct choices of f, g, and A. The most obvious ways of expressing the deblurring problem (1) in the form (5) do not lead to an efficient Douglas-Rachford iteration; in particular, for the spatially variant blur models we consider, one must not include the entire blur operator in A. Rather, the blur operator must be dissected, included partially in g and partially in A. Our method for the Nagy-O'Leary model requires that ϕ_{f} be a separable sum of functions with inexpensive proximal operators; this is a nontrivial restriction, but still allows us to handle most of the standard convex data fidelity terms, such as those using an L_1 or squared L_2 norm or the Huber penalty. For the Efficient Filter Flow model, our method requires that ϕ_{f} is the squared L_2 norm. Both methods can handle TV and tight frame regularization.

While much of the literature on image deblurring has assumed a spatially invariant blur, recent work has increasingly focused on restoring images degraded by spatially variant blur [32,2,24,22,20,23,28,3,5,4,25,15, 41]. Most of the effort has gone towards modeling and estimating spatially variant blur. In contrast, the problem of TV or tight frame regularized deblurring with a known spatially variant blur operator has not received much attention. In the original Nagy and O'Leary paper [32], non-blind deblurring is achieved by solving $Kx = b$ by the preconditioned conjugate gradient method, stopping the iteration early to mitigate the effects of noise. The same approach is taken in [2], where the Nagy-O'Leary model is used in a blind deblurring algorithm. The paper [24], which introduced the Efficient Filter Flow model, recovers x given K and b by finding a nonnegative least squares solution to $Kx = b$. (Note that [35] independently introduced a similar model.) A Bayesian framework with a hyper-Laplacian prior is used in [25] and [41], leading to non-convex optimization problems which are solved by iteratively re-weighted least squares or by adapting the method of [11]. [28] performs the non-blind deblurring step using the Richardson-Lucy algorithm. [17] models a spatially variant blur operator as being sparse in a wavelet domain and solves a standard L_2-TV deblurring problem, but does not discuss the optimization algorithm used. [20] uses the Forward-Backward algorithm to solve a spatially variant deblurring problem, but the method is restricted to the case where ϕ_{f} is the squared L_2 norm, ϕ_{c} is 0, and tight frame regularization is used. Most relevant to us, [3] uses the Nagy-O'Leary model and solves a TV deblurring problem using a method based on the domain decomposition approach given in [18], and [4] takes a similar domain decomposition approach to TV deblurring

using the Efficient Filter Flow model. However, [3] requires that ϕ_f be the squared L_2 norm, whereas our treatment of the Nagy-O'Leary model is able to handle other important data fidelity penalties such as the L_1 norm and the Huber penalty. Moreover, the papers [3,4] consider only total variation regularization, whereas our approach is able to handle tight frame regularization as well as total variation regularization. Our approach allows for constraints on x, enforced by the term $\phi_c(x)$; such constraints are not considered in [3,4]. An additional difference between our approach and that of [3,4] is that at each iteration of the methods in [3,4], there are subproblems which must themselves be solved by an iterative method; thus [3,4] present multi-level iterative algorithms, with both inner and outer iterations. In our approach all subproblems at each iteration have simple closed form solutions.

This paper is organized as follows. Section 2 briefly describes a method for solving (5) by expressing the Karush-Kuhn-Tucker (KKT) conditions as a monotone inclusion problem and applying the Douglas-Rachford splitting algorithm. (This method, as well as alternative methods such as ADMM, is discussed in greater detail in [33].) In section 3 we present the Nagy-O'Leary model of spatially variant blur and derive a method for solving (1) using this model, under the assumption that ϕ_f is separable. In section 4 we discuss the Efficient Filter Flow model of spatially variant blur, and derive a method for solving (1) using this model, in the case where ϕ_f is the squared L_2 norm. Section 5 demonstrates the effectiveness of these methods with some numerical experiments. The paper concludes with a summary of main points in section 6.

2 Primal-dual Douglas-Rachford splitting

In this section we derive the primal-dual optimality conditions for problem (5) and describe a first-order primal-dual method that we will use in this paper. We assume f and g are proper closed convex functions and A is a matrix.

2.1 Optimality conditions

To derive the optimality conditions we first reformulate (5) as

$$\begin{aligned} \underset{x,y}{\text{minimize}} \quad & f(x)+g(y) \\ \text{subject to} \quad & Ax = y. \end{aligned} \tag{6}$$

Finding primal and dual optimal variables (x,y) and z for (6) is equivalent to finding a saddle point for the Lagrangian

$$L(x,y,z) = f(x)+g(y)+\langle z,Ax\rangle - \langle z,y\rangle.$$

The variable y can be eliminated by noting that

$$\inf_{x,y} L(x,y,z) = \inf_{x} \left(f(x) - g^*(z) + \langle z,Ax\rangle\right)$$

where $g^*(z) = \sup_y(\langle z,y\rangle - g(y))$ is the convex conjugate of g. Thus the problem is reduced to finding a saddle point (x,z) of the convex-concave function

$$\ell(x,z) = f(x) - g^*(z) + \langle z,Ax\rangle.$$

The optimality conditions for this saddle point problem are

$$0 \in \partial f(x) + A^T z, \qquad 0 \in -Ax + \partial g^*(z). \tag{7}$$

Here $\partial f(x)$ and $\partial g^*(z)$ are the subdifferential of f at x and the subdifferential of g^* at z. The first condition in (7) states that x is a minimizer of $\ell(\cdot,z)$; the second condition in (7) states that z is a maximizer of $\ell(x,\cdot)$. The optimality conditions can be combined into a single condition using block notation:

$$0 \in \begin{bmatrix} 0 & A^T \\ -A & 0 \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix} + \begin{bmatrix} \partial f(x) \\ \partial g^*(z) \end{bmatrix}. \tag{8}$$

(The second term on the right denotes the Cartesian product $\partial f(x) \times \partial g^*(z)$.)

For more details on convex duality and optimality conditions, see for example [37] or [16].

2.2 Proximal operators

Before giving a method to solve the optimality condition (8), we first define and record some properties of proximal operators, a basic tool used in proximal algorithms.

Let f be a proper closed convex function on $\mathbb{R}^n$. The *proximal operator* of f is the mapping from $\mathbb{R}^n$ to $\mathbb{R}^n$ defined by

$$\mathrm{prox}_f(x) = \arg\min_u f(u) + \frac{1}{2}\|u-x\|^2.$$

It can be shown that $\mathrm{prox}_f(x)$ is uniquely defined for all x. The following facts about proximal operators will be useful to us [13, 12, 6].

1. *Moreau decomposition.* If $t > 0$, then
$$x = \mathrm{prox}_{tf}(x) + t\mathrm{prox}_{t^{-1}f^*}(x/t).$$
This rule, known as the Moreau decomposition [31], shows that the proximal operator of f^* (the convex conjugate of f) can be computed as easily as the proximal operator of f.
2. *Separable functions.* If f is separable, so that
$$f(x) = \sum_{i=1}^k f_i(x_i),$$
where x_i are subvectors of $x = (x_1, \ldots, x_k)$, then
$$\mathrm{prox}_f(x) = \begin{bmatrix} \mathrm{prox}_{f_1}(x_1) \\ \vdots \\ \mathrm{prox}_{f_k}(x_k) \end{bmatrix}. \tag{9}$$
3. *Scaling of argument.* Suppose $f(x) = g(ax)$, where $a \in \mathbb{R}$, $a \neq 0$. Then
$$\mathrm{prox}_f(x) = \frac{1}{a}\mathrm{prox}_{a^2 g}(ax). \tag{10}$$
4. *Composition with affine mapping.* Suppose $f(x) = g(Ax+b)$, where A is a matrix that satisfies $AA^T = (1/\alpha)I$ for some $\alpha > 0$. Then
$$\mathrm{prox}_f(x) = (I - \alpha A^T A)x + \alpha A^T(\mathrm{prox}_{\alpha^{-1}g}(Ax+b) - b). \tag{11}$$

For future reference we also mention a few common examples, that can be proved directly from the definition or by using the properties listed above.

1. *Quadratic function.* Suppose $f(x) = \frac{1}{2}\|Ax-b\|_2^2$, where $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m$, and $t > 0$. Then
$$\begin{aligned} \mathrm{prox}_{tf}(x) &= (I + tA^T A)^{-1}(x + tA^T b) \\ &= (I - tA^T(I + tAA^T)^{-1}A)(x + tA^T b). \end{aligned} \tag{12}$$
The second equality follows from the first by the matrix inversion lemma.
2. *Indicator function.* Suppose $f(x)$ is the indicator function of a closed convex set C and $t > 0$. Then $\mathrm{prox}_{tf}(x)$ is the Euclidean projection $\Pi_C(x)$ of x on C.
3. *Norm.* Suppose $f(x)$ is a norm and $t > 0$. Then
$$\mathrm{prox}_{tf}(x) = x - \Pi_{tC}(x) \tag{13}$$
where C is the unit ball of the dual norm. For $f(x) = \|x\|_1$, we have $tC = \{x \mid -t\mathbf{1} \leq x \leq t\mathbf{1}\}$ and (13) reduces to *soft-thresholding:*
$$\mathrm{prox}_{tf}(x)_k = \begin{cases} x_k - t & x_k > t \\ 0 & -t \leq x_k \leq t \\ x_k + t & x_k \leq -t. \end{cases}$$
If $f(x) = \|x\|$ is the Euclidean norm, then Π_{tC} is projection on a ball with radius t. If $f(x)$ is the 'isotropic' norm defined in (2), then $(u,v) = \mathrm{prox}_{tf}(x,y)$ can be computed as $(u_k, v_k) = \alpha_k(x_k, y_k)$ for $k = 1, \ldots, n$, where
$$\alpha_k = \begin{cases} 1 - \dfrac{t}{(x_k^2 + y_k^2)^{1/2}} & \text{if } (x_k^2 + y_k^2)^{1/2} > t, \\ 0 & \text{otherwise.} \end{cases}$$

2.3 Douglas-Rachford method

Problem (8) is a monotone inclusion problem and can be solved using the Douglas-Rachford splitting algorithm [7, Remark 2.9] [33]. The Douglas-Rachford method is a general algorithm for finding zeros of sums of monotone operators [29]. Applied to (8) it reduces to the following iteration:

$$\begin{aligned} x^k &= \mathrm{prox}_{tf}(p^{k-1}) \\ z^k &= \mathrm{prox}_{sg^*}(q^{k-1}) \\ \begin{bmatrix} u^k \\ v^k \end{bmatrix} &= \begin{bmatrix} I & tA^T \\ -sA & I \end{bmatrix}^{-1} \begin{bmatrix} 2x^k - p^{k-1} \\ 2z^k - q^{k-1} \end{bmatrix} \\ p^k &= p^{k-1} + \rho(u^k - x^k) \\ q^k &= q^{k-1} + \rho(v^k - z^k). \end{aligned} \tag{14}$$

We note the following useful expressions for solving the linear system in step (14):

$$\begin{aligned} \begin{bmatrix} I & tA^T \\ -sA & I \end{bmatrix}^{-1} &= \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix} + \begin{bmatrix} I \\ sA \end{bmatrix} (I + stA^TA)^{-1} \begin{bmatrix} I \\ -tA \end{bmatrix}^T \\ &= \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} -tA^T \\ I \end{bmatrix} (I + stAA^T)^{-1} \begin{bmatrix} sA^T \\ I \end{bmatrix}^T. \end{aligned}$$

There are three algorithm parameters: two step sizes $s > 0$ and $t > 0$, and a relaxation parameter $\rho \in (0,2)$. The general Douglas-Rachford iteration for solving monotone inclusion problems has only one step size. As discussed in [33], the iteration given here with two step sizes can be derived by using Douglas-Rachford (with a single step size t) to solve the optimality condition (8) for the modified problem

$$\underset{x}{\text{minimize}} \quad f(x) + \tilde{g}(\tilde{A}x) \tag{15}$$

where $\tilde{A} = \beta A$ and $\tilde{g}(y) = g(y/\beta)$ with $\beta = \sqrt{s/t}$.

As can be seen from the algorithm outline, the primal-dual Douglas-Rachford method is particularly useful for problems that satisfy the following two assumptions.

A1. f and g^* have inexpensive proximal operators.

A2. The linear system in step (14) can be solved efficiently by exploiting structure in A. More specifically, linear systems of the form $(I + stA^TA)x = y$ can be solved efficiently.

Due to the necessity of satisfying these two conditions, utilizing the Douglas-Rachford method is not always a straightforward exercise. In any given application, one must find a suitable choice of f, g, and A, which is not always possible. A more detailed discussion of this method, as well as other methods for solving (5) under these assumptions, can be found in [33,34,26].

Although in this paper we focus on the primal-dual Douglas-Rachford splitting method we should point out that the techniques of sections 3 and 4 can also be used in conjunction with ADMM [19,10]. To apply ADMM, we first reformulate problem (5) as

$$\begin{aligned} \underset{x,y,u}{\text{minimize}} \quad & f(u) + g(y) \\ \text{subject to} \quad & u = x \\ & y = Ax, \end{aligned}$$

where we have introduced a "splitting variable" u. This reformulated problem can be solved efficiently by ADMM, but only when f, g, and A are chosen so that assumptions A1 and A2 are satisfied.

3 The Nagy-O'Leary model

The Nagy-O'Leary model of spatially variant blur [32] assumes that the blur operator K has the form

$$K = \sum_{p=1}^{P} U_p K_p. \tag{16}$$

where

- each K_p, for $p = 1, \ldots, P$, performs a convolution with a space-invariant kernel,
- each U_p is diagonal with nonnegative diagonal entries,
- the matrices U_p sum to the identity matrix: $\sum_{p=1}^{P} U_p = I$.

Intuitively, U_p determines how much K_p contributes to each pixel in the blurry image.

We now present a method for solving problem (1) when K has the form (16). The method requires the special assumption that ϕ_{f} is a separable sum of functions with inexpensive proximal operators:

$$\phi_{\mathrm{f}}(x) = \sum_i \phi_{\mathrm{f}}^i(x_i).$$

This is the case in most applications, for example if ϕ_{f} is a squared L_2 norm ($\phi_{\mathrm{f}}^i(u) = u^2/2$), an L_1 norm ($\phi_{\mathrm{f}}^i(u) = |u|$), or the Huber penalty

$$\phi_{\mathrm{f}}^i(u) = \begin{cases} u^2/(2\eta) & |u| \le \eta \\ |u| - \eta/2 & |u| \ge \eta, \end{cases}$$

where η is a positive parameter. We make no assumptions on ϕ_{r} and ϕ_{c}, except that they have inexpensive proximal operators. The method also requires that $D^T D$ is diagonalized by the discrete Fourier basis, which is true for both TV and tight frame regularization. Assuming K satisfies (16), we write problem (1) in the generic form (5) by defining

$$\begin{aligned} f(x) &= \phi_{\mathrm{c}}(x), \\ g(y_1, \ldots, y_P, w) &= \phi_{\mathrm{f}}(U_1 y_1 + \cdots + U_P y_P - b) + \phi_{\mathrm{r}}(w), \\ A &= \begin{bmatrix} K_1^T & \cdots & K_P^T & D^T \end{bmatrix}^T. \end{aligned}$$

Note that the blur operator K has been dissected, included partially in A and partially in g. We can apply the primal-dual splitting method of section 2 to this problem, once we first check that the assumptions A1 and A2 of section 2 are satisfied.

We first examine assumption A1. The proximal operator of f is the proximal operator of ϕ_{c}, which is assumed to be inexpensive. The function g is separable in the variables $y = (y_1, \ldots, y_P)$ and w, and can be written as $g(y,w) = g_1(y) + g_2(w)$ where

$$g_1(y) = \phi_{\mathrm{f}}(U_1 y_1 + \cdots + U_P y_P - b), \qquad g_2(w) = \phi_{\mathrm{r}}(w).$$

Therefore the proximal operator of g at $(\hat{y}, \hat{w})$ is of the form $\mathrm{prox}_{tg}(\hat{y}, \hat{w}) = (\mathrm{prox}_{tg_1}(\hat{y}), \mathrm{prox}_{tg_2}(\hat{w}))$. The proximal operator of g_2 is the proximal operator of ϕ_{r}, which is assumed to be inexpensive.

It's not yet obvious that the proximal operator of g_1 can be calculated efficiently. Indeed, if ϕ_{f} were not assumed to be separable, then this would not be the case. Let $U = \begin{bmatrix} U_1 & \cdots & U_P \end{bmatrix}$ and let $M = (UU^T)^{1/2}$. Then we can express g_1 as $g_1(y) = h(M^{-1}Uy - M^{-1}b)$, where

$$h(u) = \phi_{\mathrm{f}}(Mu) = \sum_i \phi_{\mathrm{f}}^i(M_{ii} u_i).$$

The invertibility of M follows from $\sum_{p=1}^{P} U_p = I$. Noting that $(M^{-1}U)(M^{-1}U)^T = I$, we can now use the composition rule (11) to express the proximal operator of g_1 in terms of the proximal operator of h. Moreover, the proximal operator of h can be evaluated efficiently using the separable sum rule (9) together with the rule (10). The resulting formula for the proximal operator of g_1 is

$$\mathrm{prox}_{tg_1}(\hat{y}) = (I - U^T M^{-2} U)\hat{y} + U^T M^{-2}(v + b), \tag{17}$$

where v is the vector whose ith component is

$$v_i = \mathrm{prox}_{tM_{ii}^2 \phi_{\mathrm{f}}^i}((U\hat{y} - b)_i). \tag{18}$$

Regarding A2, note that

$$A^T A = K_1^T K_1 + \cdots + K_P^T K_P + D^T D.$$

If D represents a discrete gradient operator using periodic boundary conditions, or if D represents a tight frame analysis operator (*i.e.*, $D^T D = \alpha I$) then $D^T D$ is diagonalized by the discrete Fourier basis. In this case, $A^T A$ is also diagonalized by the discrete Fourier basis, and linear systems with coefficient $I + \lambda A^T A$ can be solved efficiently via the FFT.

It's physically unrealistic to assume periodic boundary conditions when modeling the blurring process. On the other hand, we want the operators K_p to use periodic boundary conditions for computational efficiency. We can deal with this issue by using an "unknown boundary conditions" approach as in [1]. Recall that we're assuming ϕ_{f} is a separable sum:

$$\phi_{\mathrm{f}}(x) = \sum_i \phi_{\mathrm{f}}^i(x_i).$$

If pixel i is a border pixel (close enough to the edge of the image that boundary conditions are relevant), then we simply redefine the function ϕ_{f}^i to be identically zero. We then solve the same optimization problem as before. The fact that the operators K_p use periodic boundary conditions is no longer a problem. In this approach, it is as if we are "inpainting" the border of the recovered image x. Once we've computed x, we can then discard the inpainted border. (Discarding the border also removes any artifacts introduced by using periodic boundary conditions for D.)

4 The Efficient Filter Flow model

A popular variant of the Nagy-O'Leary model for spatially variant blur is the *Efficient filter flow* (EFF) model introduced in [24]. This model is used in the recent papers [22,23,5]. In the EFF model the blur operator K has the form

$$K = \sum_{p=1}^{P} K_p U_p \tag{19}$$

where the matrices K_p and U_p satisfy the same properties as in the Nagy-O'Leary model of section 3. In this section, we present a method for solving

$$\underset{x}{\text{minimize}} \quad \frac{1}{2}\|Kx - b\|^2 + \phi_{\mathrm{r}}(Dx) + \phi_{\mathrm{c}}(x) \tag{20}$$

in the case where K is given by (19). Problem (20) is the special case of problem (1) where $\phi_{\mathrm{f}}(x) = \frac{1}{2}\|x\|^2$. We assume, as before, that ϕ_{r} and ϕ_{c} are proper closed convex functions with inexpensive proximal operators. Additionally, we assume that D is a tight frame operator ($D^T D = \alpha I$), but the method can also be used when $D^T D$ is very sparse (as in TV regularization).

Problem (20) can be expressed in the generic form (5) by defining

$$f(x) = \phi_{\mathrm{c}}(x), \tag{21a}$$

$$g(y_1, \ldots, y_P, w) = \frac{1}{2}\|\sum_{p=1}^{P} K_p y_p - b\|_2^2 + \phi_{\mathrm{r}}(w), \tag{21b}$$

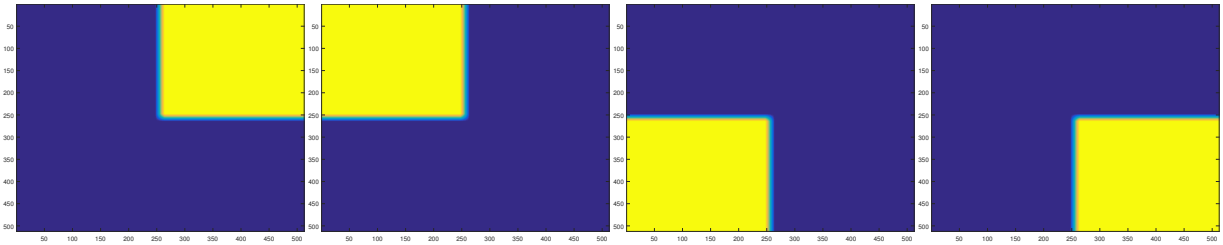
$$A = \begin{bmatrix} U_1^T & \cdots & U_P^T & D^T \end{bmatrix}^T. \tag{21c}$$

We can solve this problem using the primal-dual splitting method of section 2, but we first must check that the assumptions A1 and A2 of section 2 are satisfied.

Again we first examine assumption A1. The proximal operator of f is the proximal operator of ϕ_{c}, which is assumed to be inexpensive. The function g is separable in the variables $y = (y_1, \ldots, y_P)$ and w, and can be written as $g(y, w) = g_1(y) + g_2(w)$ where

$$g_1(y) = \frac{1}{2}\|K_1 y_1 + \cdots + K_P y_P - b\|_2^2, \qquad g_2(w) = \phi_{\mathrm{r}}(w).$$

Therefore the proximal operator of g at $(\hat{y}, \hat{w})$ is of the form $\mathrm{prox}_{tg}(\hat{y}, \hat{w}) = (\mathrm{prox}_{tg_1}(\hat{y}), \mathrm{prox}_{tg_2}(\hat{w}))$. The proximal operator of g_2 is the proximal operator of ϕ_{r}, which is assumed to be inexpensive. The key point of this section is that g_1 also has an inexpensive proximal operator. Once we show this, it will follow that assumption A1 is satisfied.



(a) U_1. (b) U_2 (c) U_3 (d) U_4

Fig. 1 Visualization of the matrices U_p in section 5.1. Red corresponds to a value of 1, and blue corresponds to a value of 0. The numerical values transition smoothly from 0 to 1.

We can express g_1 as $g_1(y) = (1/2)\|By-b\|_2^2$, where $B = \begin{bmatrix} K_1 & \cdots & K_P \end{bmatrix}$. From rule (12) for the proximal operator of a quadratic function, the proximal operator of g_1 is given by

$$\text{prox}_{tg_1}(\hat{y}) = (I - tB^T(I + tBB^T)^{-1}B)(\hat{y} + tB^T b). \tag{22}$$

Because the matrix

$$I + tBB^T = I + t\sum_{p=1}^{P} K_p K_p^T$$

is diagonalized by the discrete Fourier basis, expression (22) for the proximal operator of g_1 can be evaluated efficiently using the FFT.

Regarding A2, note that the linear equation in each iteration has coefficient matrix

$$I + stA^T A = I + st(U_1^2 + \cdots + U_P^2 + D^T D).$$

This matrix is diagonal when D represents the analysis operator of a tight frame, and very sparse when D represents a discrete gradient operator.

5 Experiments

All experiments were performed on a computer with a 3.70 GHz Intel Xeon(R) E5-1620 v2 processor with 8 cores and 7.7 GB of RAM. The code was written in MATLAB using MATLAB version 8.1.0.604 (R2013a).

5.1 Spatially variant Gaussian blur

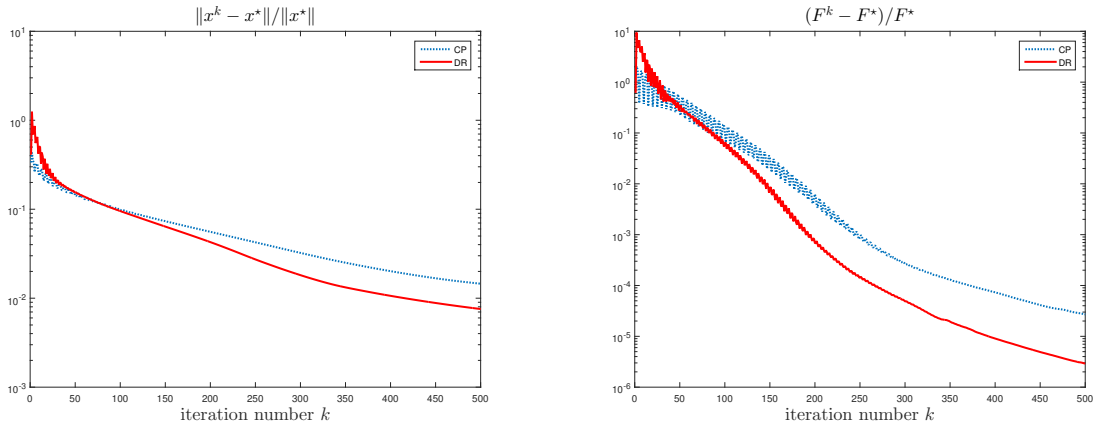
We demonstrate the method of section 3 by restoring an image which has been blurred by an operator K that is described by the Nagy-O'Leary model (16). The image used is the 512 by 512 "Barbara" image, scaled to have intensity values between 0 and 1. We set $P = 4$, corresponding to the four quadrants of the image. For $p = 1, \ldots, 4$, the matrix K_p performs a convolution with a 17 by 17 truncated Gaussian kernel, with standard deviation $\sigma_p = p$ pixels. The matrix U_p zeros out components away from the pth quadrant, as visualized in figure 1 (where red corresponds to a value of 1 and blue corresponds to a value of 0). When creating the blurry image, "replicate" boundary conditions were used – meaning that an intensity value at a location outside the image is assumed to be equal to the intensity at the nearest pixel in the image. Gaussian noise with zero mean and standard deviation 10^{-3} was added to each pixel of the blurred image. After Gaussian noise was added to each pixel, 10% of the pixels (chosen at random) were corrupted by "salt and pepper" noise. (The intensity value at each corrupted pixel was randomly set to either 0 or 1.)

We handle boundary conditions as described in section 3, using the "unknown boundary conditions" approach. First the blurry image is zero padded to have size 528 by 528, then a restored image is computed by solving

$$\underset{x}{\text{minimize}} \quad \phi_{\text{f}}(Kx - b) + \gamma\|Dx\|_{\text{iso}}. \tag{23}$$

(a) Original image. (b) Blurry, noisy image. (c) Restored image.

Fig. 2 Result for the experiment in section 5.1.



(a) Relative error vs. iteration. (b) Relative optimality gap vs. iteration.

Fig. 3 Relative error vs. iteration and relative optimality gap vs. iteration for the experiment in section 5.1. The solid line shows the convergence of the primal-dual Douglas-Rachford method, and the dashed line shows the convergence of the Chambolle-Pock method.

We take $\phi_{\mathrm{f}}(x) = \sum_i \phi_{\mathrm{f}}^i(x_i)$, where ϕ_{f}^i is the Huber penalty (with parameter $\eta = 10^{-3}$) if pixel i is not a border pixel, and ϕ_{f}^i is identically zero if pixel i is a border pixel. The optimization variable is $x \in \mathbb{R}^{528^2}$, and D is the matrix representation of a discrete gradient operator $\tilde{D} : \mathbb{R}^{528\times 528} \to \mathbb{R}^{528\times 528\times 2}$, defined by:

$$(\tilde{D}x)_{i,j,1} = x_{i+1,j} - x_{ij}, \qquad (\tilde{D}x)_{i,j,2} = x_{i,j+1} - x_{ij}. \tag{24}$$

($\tilde{D}$ uses periodic boundary conditions.) Once a deblurred image is computed by solving (23), the inpainted boundary is discarded, yielding a restored image of size 512 by 512. The parameter γ is chosen to give a visually appealing image reconstruction. The original, blurry, and restored images are shown in figure 2. Notice that the lower right quadrant is blurred the most. It's interesting to look closely and see that more detail is recovered in the upper right quadrant, where the blurring was less severe.

In figure 3a, the quantity $\|x^k - x^\star\|/\|x^\star\|$ is plotted against the iteration number k. Here x^k is the estimate of the solution to (23) at iteration k, and $x^\star$ is a nearly optimal primal variable which was computed by running the method of section 3 for 10,000 iterations. In figure 3b, the quantity $(F^k - F^\star)/F^\star$ is plotted against the iteration number k. Here F^k is the primal objective function value at iteration k, and $F^\star$ is a nearly optimal primal objective function value, which was computed by running the method of section 3 for 10,000 iterations. To illustrate the convergence properties of the algorithm, we show the error for the first 500 iterations. However we observed that after about 180 iterations the estimate x^k was visually indistinguishable from $x^\star$.

Figure 3 compares the performance of the method of section 3 (labeled "DR") with the performance of the well known Chambolle-Pock method [9] (labeled "CP"). Chambolle-Pock minimizes $f(x) + g(Ax)$, where f and g are

proper closed convex functions, via the iteration

$$\begin{aligned}\bar{x}^k &= \mathrm{prox}_{tf}(x^{k-1} - tA^T z^{k-1}),\\ \bar{z}^k &= \mathrm{prox}_{sg^*}(z^{k-1} + sA(2\bar{x}^k - x^{k-1})),\\ (x^k, z^k) &= \rho(\bar{x}^k, \bar{z}^k) + (1-\rho)(x^{k-1}, z^{k-1}).\end{aligned}$$

Here s and t are step sizes, and $\rho \in (0,2)$ is an overrelaxation parameter. (This overrelaxed version of Chambolle-Pock is presented in [14].) The step sizes s and t are required by convergence proofs to satisfy $st\|A\|^2 \leq 1$ [14], and we choose them so that $st\|A\|^2 = 1$. We precompute $\|A\|$ using power iteration. When solving (23) by Chambolle-Pock, we take

$$A = \begin{bmatrix} K \\ D \end{bmatrix}$$

and

$$f(x) = 0, \qquad g(y_1, y_2) = \phi_{\mathrm{f}}(y_1 - b) + \gamma\|y_2\|_{\mathrm{iso}}$$

for all x, y_1, y_2. One of the main advantages of Chambolle-Pock is that it requires only *applying* A and A^T (at each iteration), and never requires solving a linear system involving A. It is therefore very well suited to exploit the structure in the Nagy-O'Leary model.

As shown in figure 3, in this example the method of section 3 converges in fewer iterations than the Chambolle-Pock method. Moreover, the time per iteration is nearly the same for both methods: 0.19 seconds for Chambolle-Pock compared to 0.20 seconds for the method of section 3. In this experiment, close to optimal fixed step sizes and overrelaxation parameters for both methods were chosen by trial and error. (In particular, the Chambolle-Pock step sizes were not taken to be equal.)

5.2 Motion deblurring

In this section we combine our approach to spatially variant deblurring with the motion segmentation algorithm of [8]. The algorithm presented in [8] segments out a motion-blurred region from an otherwise sharp input image and estimates a motion blur kernel for this region, but does not compute a deblurred image.

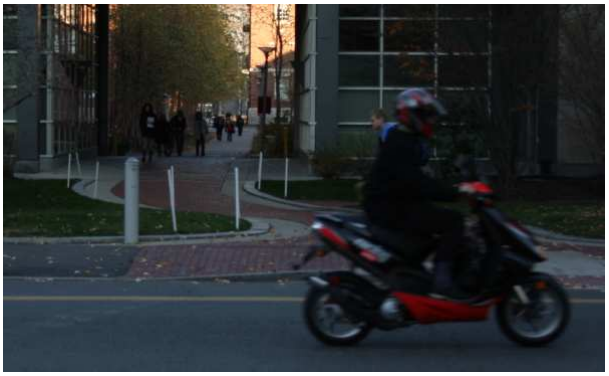
A blurry image (of size 367 by 600) is shown in figure 4a, and a segmentation of this image computed using the code provided by [8] is shown in figure 5a.

The algorithm of [8] estimated a horizontal motion of 6 pixels during the time the camera shutter was open, but a slightly better image reconstruction was obtained using the motion blur kernel $(1/7)\begin{bmatrix}1 & 1 & 1 & 1 & 1 & 1 & 1\end{bmatrix}$, which corresponds to a horizontal motion of 7 pixels.

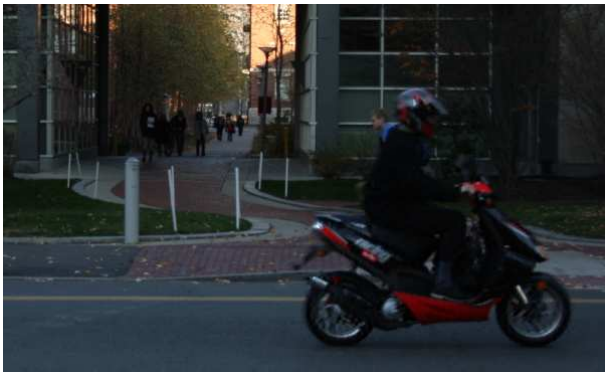
We use the Nagy-O'Leary model (16) with $P = 2$. The matrix K_1 performs a convolution with the 7 pixel motion blur kernel given above, and K_2 is the identity matrix (we assume negligible blur for the background region). The matrix U_1 zeros out background pixels without altering foreground intensity values, while the matrix U_2 zeros out foreground pixels without altering background intensity values. We first zero pad the blurry image to have size 373 by 606, then compute a deblurred image by solving (23) using the method of section 3. (We deblur each color channel of the blurry image separately, to obtain a restored color image.) We use a quadratic data fidelity function ϕ_{f}. The discrete gradient operator D is defined as in section 5.1. The parameter γ is chosen to give a visually appealing image reconstruction. Once a deblurred image is computed by solving (23), the inpainted boundary is discarded, yielding a restored image of size 367 by 600. (Because K_2 is the identity, this effort to handle boundary conditions correctly is actually unnecessary in this example.)

When solving (23), we ran the primal-dual Douglas-Rachford method discussed in section 2 for 250 iterations. We also solved problem (23) using the Chambolle-Pock algorithm, as described in section 5.1, but taking $g(y_1, y_2) = (1/2)\|y_1 - b\|^2 + \gamma\|y_2\|_{\mathrm{iso}}$. The step sizes and overrelaxation parameters for both methods were set to the same values used in section 5.1. In figure 5b, the quantity $(F^k - F^\star)/F^\star$ is plotted against the iteration number k. Here F^k is the primal objective function value at iteration k, and $F^\star$ is a nearly optimal primal objective function value, which was computed by running the method of section 3 for $10,000$ iterations. The time per iteration was 0.23 seconds for both methods.

The deblurred image is shown in figure 4b. When we zoom in, the letters on the motorcycle in the restored image are nearly legible now, and appear to say "racing".



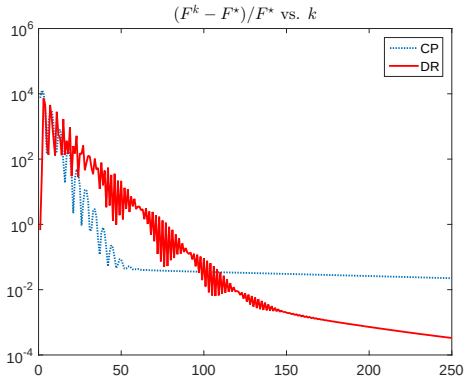
(a) Blurry image.



(b) Restored image.

Fig. 4 Result for the experiment in section 5.2.

(a) Segmentation.



(b) Convergence plot.

Fig. 5 Segmentation and convergence plot for the experiment in section 5.2. The solid line shows the convergence of the primal-dual Douglas-Rachford method, and the dashed line shows the convergence of the Chambolle-Pock algorithm.

5.3 Blur due to camera shake

In this section we use the method of section 4, based on the Efficient Filter Flow model, to deblur an 800 by 800 color image which is blurry due to camera shake. The image is taken from the benchmark dataset [42] which provides ground truth spatially variant blur kernels for real-world blurred images. (The benchmark dataset blur kernels were obtained by precisely measuring the camera trajectory and rotation as the images were captured.) We use the Efficient Filter Flow model (19) with $P = 16$, partitioning an image into a 4 by 4 grid of subimages. The matrix U_p zeros out components away from the pth subimage, as in section 5.1 (but now with $P = 16$). The blurry image is scaled so that RGB values are between 0 and 1 and each color channel is deblurred separately by solving problem (20). The penalty function ϕ_r in equation (20) is taken to be the L_1-norm, ϕ_c is identically zero, and D is the analysis operator for a curvelet tight frame [43]. The parameter γ was chosen to give a visually appealing image reconstruction. Because our method for the Efficient Filter Flow model assumes periodic boundary conditions, each color channel $b \in \mathbb{R}^{800\times 800}$ in the blurry image is first extended to an (approximately) periodic image $\tilde{b} \in \mathbb{R}^{816\times 816}$ by solving

$$\begin{array}{ll} \underset{x\in\mathbb{R}^{816\times 816}}{\text{minimize}} & \frac{1}{2}\|\tilde{D}x\|^2 \\ \text{subject to} & x_{i+8,j+8} = b_{i,j} \quad \text{for all } 1 \le i \le 800, 1 \le j \le 800, \end{array} \tag{25}$$

where $\tilde{D}$ is a discrete gradient operator using periodic boundary conditions (see equation (24)). Problem (25) is a standard linear algebra problem that can be solved extremely efficiently due to the fact that $\tilde{D}$ is diagonalized by the discrete Fourier basis (so linear systems of equations involving $\tilde{D}$ can be solved using the fast Fourier transform). Once a deblurred image is computed by solving (20), the artificial boundary is discarded, yielding a restored image of size 800 by 800.

(a) Blurry image. (b) Restored image.

Fig. 6 Result for the experiment in section 5.3.

For comparison, we solve the same deblurring problem using the Chambolle-Pock algorithm. When solving (20) with the Chambolle-Pock algorithm, a standard approach would be to take

$$A = \begin{bmatrix} K \\ D \end{bmatrix}$$

and

$$f(x) = 0, \qquad g(y_1, y_2) = \frac{1}{2}\|y_1 - b\|^2 + \gamma\|y_2\|_1$$

for all x, y_1, y_2. Here K is given by equation (19). To satisfy the Chambolle-Pock step size restriction, we require an upper bound on the norm of A. In the standard approach, the norm of A can be estimated using $\|A\| \leq \sqrt{\|K\|^2 + \|D\|^2}$, and the norm of K can be estimated using

$$\|K\| \leq \sum_{p=1}^{P} \|K_p\| \|U_p\|. \tag{26}$$

The norms of the matrices K_p and U_p are known exactly because each K_p is diagonalized by the discrete Fourier basis, and each U_p is diagonal. This bound on $\|A\|$ might be overly conservative, however, which can lead to slow convergence of the Chambolle-Pock algorithm. This is particularly a problem for large values of P.

A better way to apply the Chambolle-Pock algorithm is to make use of the dissection of K which is presented in section 4. In this "dissection" approach, we take A, f, and g as in equations (21). The prox-operator of g can be evaluated efficiently using the method presented in section 4. The dissection approach has the advantage that the norm of A is known exactly, because $A^T A = \sum_{p=1}^{P} U_p^T U_p + D^T D$ is diagonal. (Note that $D^T D = \lambda I$, because D is a tight frame analysis operator. For the curvelet tight frame used in this experiment, $\lambda = 1$.)

The blurry and deblurred images are shown in figure 6. Figure 7 shows plots of objective function value versus iteration for the method presented in section 4 (using the Douglas-Rachford algorithm), as well as for the Chambolle-Pock algorithm (both the standard and the "dissection" approaches). The time per iteration was 3.23 seconds for the standard Chambolle-Pock implementation, 3.33 seconds for Chambolle-Pock with dissection, and 3.36 seconds for Douglas-Rachford. The step sizes s and t are chosen to satisfy the Chambolle-Pock step size restriction $st\|A\|^2 \leq 1$. In this experiment, close to optimal fixed step sizes and overrelaxation parameters for both methods were chosen by trial and error. In particular, the Chambolle-Pock step sizes were not taken to be equal.

An advantage of the Douglas-Rachford method is that there is no need to estimate the norm of the matrix A. As seen in figure 7, this can lead to faster convergence than a standard implementation of the Chambolle-Pock algorithm using the bound (26). Figure 7 also shows that the dissection of K presented in section 4 allows for an improved Chambolle-Pock implementation, with faster convergence because the norm of A is known exactly. There is no need to use an expensive iterative method such as power iteration to compute the norm of A.

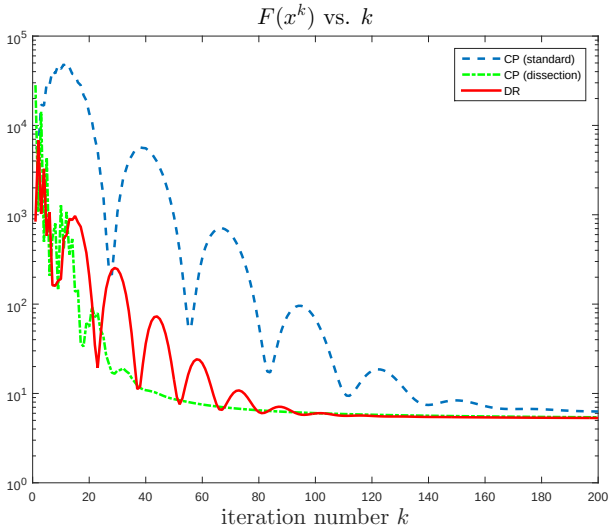


Fig. 7 Objective function value versus iteration for the experiment in section 5.3. The solid line shows the convergence of the primal-dual Douglas-Rachford method, and the dashed lines show the convergence of the Chambolle-Pock method both with and without operator dissection.

6 Conclusion

Beginning with [38] over 20 years ago, a substantial literature has been devoted to non-blind TV image deblurring under the assumption of spatially invariant blur. In this paper, we have extended this line of research by presenting efficient methods for TV or tight frame regularized deblurring using two fundamental models of *spatially variant* blur: the classical Nagy-O'Leary model and the related Efficient Filter Flow model. In the case of the Nagy-O'Leary model, our method requires that the data fidelity function ϕ_{f} in problem (1) is a separable sum of functions with inexpensive proximal operators. This includes most standard data fidelity functions such as the squared L_2 norm, the L_1 norm, and the Huber penalty. In the case of the Efficient Filter Flow model, our method requires that ϕ_{f} is the squared L_2 norm. Both methods can handle TV and tight frame regularization, as well as constraints such as box constraints on the recovered image x.

The Nagy-O'Leary model and the Efficient Filter Flow model both express a spatially variant blur operator as a sum of P terms, each term involving a spatially invariant blur operator. For the non-blind deblurring algorithms presented in this paper, the computational cost is dominated by the cost of computing order P fast Fourier transforms at each iteration. The cost per iteration is $O(PN^2 \log N)$ for N by N images. Thus, we can solve (1), for the two spatially variant blur models, with an efficiency that is comparable with the efficiency of FFT-based methods that assume a spatially invariant blur model. In the case where $P = 1$, the methods presented in this paper reduce to state of the art proximal algorithms for spatially invariant TV or tight frame regularized deblurring (see [33] for example).

A recurring theme of research into algorithms for large scale convex optimization has been the discovery that in many applications, with appropriate splitting techniques, implicit Douglas-Rachford-based methods (including ADMM) can be implemented with the same cost per iteration as simpler methods such as the Chambolle-Pock algorithm. We have shown this to be the case for spatially variant TV or tight frame deblurring with the Nagy-O'Leary and Efficient Filter Flow blur models. Douglas-Rachford-based methods have the advantage that there is no step size restriction, and so it is unnecessary to estimate operator norms. Moreover, in some cases it may be easier to achieve high accuracy using the Douglas-Rachford method, as suggested by the experiments in sections 5.1 and 5.2. As demonstrated in the experiment in section 5.3, the operator "dissections" presented in sections 3 and 4 are useful even when using the Chambolle-Pock algorithm, because with this approach exact analytic formulas are available for the operator norms appearing in the Chambolle-Pock step size restriction.

The experiment in section 5.2 gives an example where the methods of this paper can be combined with a motion segmentation algorithm to obtain a blind motion deblurring algorithm. In future work, it would be interesting to explore combining the non-blind algorithms presented here with successful approaches to blind deblurring, such as recent algorithms that restore images degraded by camera shake. Along these lines, the paper [23] presents a blind deblurring algorithm that combines the structural constraints of a Projective Motion Path Blur (PMPB) model with the efficiency of the Efficient Filter Flow model, reaping the benefits of both frameworks. This demonstrates

that PMPB models, a subject of much current interest, can be usefully combined with classical models of spatially variant blur, where our methods are applicable. Another goal would be to develop proximal algorithms that work with PMPB models directly.

Many blind deblurring algorithms, including [23], use a non-convex regularizer based on the statistics of natural images, rather than using TV regularization. Another future research direction is to adapt the proximal algorithms given in this paper, to handle these non-convex regularizers.

References

1. M. S. C. Almeida and M. A. T. Figueiredo. Deconvolving images with unknown boundaries using the alternating direction method of multipliers. *IEEE Trans. Image Process.*, 22(8):3074–3086, 2013.
2. J. Bardsley, S. Jefferies, J. Nagy, and R. Plemmons. A computational method for the restoration of images with an unknown, spatially-varying blur. *Optics express*, 14(5):1767–1782, 2006.
3. S. Ben Hadj and L. Blanc Féraud. Restoration method for spatially variant blurred images. Rapport de recherche RR-7654, INRIA, June 2011.
4. S. Ben Hadj and L. Blanc-Féraud. Modeling and removing depth variant blur in 3d fluorescence microscopy. In *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, pages 689–692. IEEE, 2012.
5. S. Ben Hadj, L. Blanc-Féraud, G. Aubert, et al. Space variant blind image restoration. 2012.
6. S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found. Trends Mach. Learn.*, 3(1):1–122, January 2011.
7. L. M. Briceno-Arias and P. L. Combettes. A monotone+ skew splitting model for composite monotone inclusions in duality. *SIAM Journal on Optimization*, 21(4):1230–1250, 2011.
8. A. Chakrabarti, T. Zickler, and W. T. Freeman. Analyzing spatially-varying blur. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 2512–2519. IEEE, 2010.
9. A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of Mathematical Imaging and Vision*, 40:120–145, 2011.
10. R. H. Chan, M. Tao, and X. Yuan. Constrained total variation deblurring methods and fast algorithms based on alternating direction method of multipliers. *SIAM Journal on Imaging Sciences*, 6(1):680–697, 2013.
11. S. Cho and S. Lee. Fast motion deblurring. *ACM Trans. Graph.*, 28(5):145:1–145:8, December 2009.
12. P. L. Combettes and J.-C. Pesquet. A Douglas-Rachford splitting approach to nonsmooth convex variational signal recovery. *IEEE Journal of Selected Topics in Signal Processing*, 1(4):564–574, 2007.
13. P. L. Combettes and V. R. Wajs. Signal recovery by proximal forward-backward splitting. *Multiscale Modeling and Simulation*, 4(4):1168–1200, 2005.
14. L. Condat. A primal-dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms. *Journal of Optimization Theory and Applications*, 158(2):460–479, 2013.
15. L. Denis, E. Thiebaut, and F. Soulez. Fast model of space-variant blurring and its application to deconvolution in astronomy. In *Image Processing (ICIP), 2011 18th IEEE International Conference on*, pages 2817–2820, Sept 2011.
16. I. Ekeland and R. Témam. *Convex Analysis and Variational Problems*, volume 28 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics, 1999. First published in 1976 by North-Holland.
17. P. Escande, P. Weiss, and F. Malgouyres. Image restoration using sparse approximations of spatially varying blur operators in the wavelet domain. In *Journal of Physics: Conference Series*, volume 464, page 012004. IOP Publishing, 2013.
18. M. Fornasier, A. Langer, and C.-B. Schnlieb. A convergent overlapping domain decomposition method for total variation minimization. *Numerische Mathematik*, 116(4):645–685, 2010.
19. T. Goldstein and S. Osher. The split Bregman method for L1-regularized problems. *SIAM Journal on Imaging Sciences*, 2(2):323–343, 2009.
20. N. Hajlaoui, C. Chaux, G. Perrin, F. Falzon, and A. Benazza-Benyahia. Satellite image restoration in the context of a spatially varying point spread function. *J. Opt. Soc. Am. A*, 27(6):1473–1481, 2010.
21. P. C. Hansen, J. G. Nagy, and D. P. O'Leary. *Deblurring Images. Matrices, Spectra, and Filtering*. Society for Industrial and Applied Mathematics, 2006.
22. S. Harmeling, M. Hirsch, and B. Schölkopf. Space-variant single-image blind deconvolution for removing camera shake. In J.D. Lafferty, C.K.I. Williams, J. Shawe-Taylor, R.S. Zemel, and A. Culotta, editors, *Advances in Neural Information Processing Systems 23*, pages 829–837. Curran Associates, Inc., 2010.
23. M. Hirsch, C. J. Schuler, S. Harmeling, and B. Scholkopf. Fast removal of non-uniform camera shake. In *Proceedings of the 2011 International Conference on Computer Vision*, ICCV '11, pages 463–470, Washington, DC, USA, 2011. IEEE Computer Society.
24. M. Hirsch, S. Sra, B. Scholkopf, and S. Harmeling. Efficient filter flow for space-variant multiframe blind deconvolution. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 607–614, June 2010.
25. N. Joshi, S. B. Kang, C. L. Zitnick, and R. Szeliski. Image deblurring using inertial measurement sensors. *ACM Trans. Graph.*, 29(4):30:1–30:9, July 2010.
26. N. Komodakis and J.-C. Pesquet. Playing with duality: An overview of recent primal-dual approaches for solving large-scale optimization problems. *arXiv preprint arXiv:1406.5429*, 2014.
27. G. Kutyniok, M. Shahram, and X. Zhuang. Shearlab: A rational design of a digital parabolic scaling algorithm. *arXiv preprint arXiv:1106.1319*, 2011.
28. A. Levin. Blind motion deblurring using image statistics. In *NIPS*, volume 2, page 4, 2006.
29. P. L. Lions and B. Mercier. Splitting algorithms for the sum of two nonlinear operators. *SIAM Journal on Numerical Analysis*, 16(6):964–979, 1979.
30. S. Mallat. *A Wavelet Tour of Signal Processing*. Academic Press, second edition, 1999.
31. J. J. Moreau. Proximité et dualité dans un espace hilbertien. *Bull. Math. Soc. France*, 93:273–299, 1965.
32. J. G. Nagy and D. P. O'Leary. Restoring images degraded by spatially variant blur. *SIAM J. Sci. Comput.*, 19(4):1063–1082, July 1998.

33. D. O'Connor and L. Vandenberghe. Primal-dual decomposition by operator splitting and applications to image deblurring. *SIAM Journal on Imaging Sciences*, 7:1724–1754, 2014.
34. Patrick L P. L. Combettes and J.-C. Pesquet. Primal-dual splitting algorithm for solving inclusions with mixtures of composite, lipschitzian, and parallel-sum type monotone operators. *Set-Valued and variational analysis*, 20(2):307–330, 2012.
35. C. Preza and J.-A. Conchello. Depth-variant maximum-likelihood restoration for three-dimensional fluorescence microscopy. *JOSA A*, 21(9):1593–1601, 2004.
36. N. Pustelnik, C. Chaux, and J.-C. Pesquet. Parallel proximal algorithm for image restoration using hybrid regularization. *IEEE Trans. Image Process.*, 20(9):2450–2462, 2011.
37. R. T. Rockafellar. *Convex Analysis*. Princeton Univ. Press, Princeton, second edition, 1970.
38. L. Rudin, S. J. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.
39. L.I. Rudin and S. Osher. Total variation based image restoration with free local constraints. In *Image Processing, 1994. Proceedings. ICIP-94., IEEE International Conference*, volume 1, pages 31–35 vol.1, Nov 1994.
40. C. R. Vogel. *Computational Methods for Inverse Problems*. Society for Industrial and Applied Mathematics, 2002.
41. O. Whyte, J. Sivic, A. Zisserman, and J. Ponce. Non-uniform deblurring for shaken images. *International Journal of Computer Vision*, 98(2):168–186, 2012.
42. R. Köhler, M. Hirsch, B. Mohler, B. Schölkopf, and S. Harmeling. Recording and playback of camera shake: Benchmarking blind deconvolution with a real-world database. In *Computer Vision–ECCV 2012*, pages 27–40. Springer, 2012.
43. E. Candes, L. Demanet, D. Donoho, and L. Ying. Fast discrete curvelet transforms. *Multiscale Modeling & Simulation*, 5(3):861–899, 2006.
44. T. Pock, D. Cremers, H. Bischof, and A. Chambolle. An algorithm for minimizing the mumford-shah functional. In *Computer Vision, 2009 IEEE 12th International Conference on*, pages 1133–1140. IEEE, 2009.