PAPER *Special Section on Picture Coding and Image Media Processing*

Image Restoration with Multiple Hard Constraints on Data-Fidelity to Blurred/Noisy Image Pair

Saori TAKEYAMA[†a)], ***Nonmember***, **Shunsuke ONO**[††,†††], ***Member***, ***and*** **Itsuo KUMAZAWA**[††,†††], ***Nonmember***

SUMMARY Existing image deblurring methods with a blurred/noisy image pair take a two-step approach: blur kernel estimation and image restoration. They can achieve better and much more stable blur kernel estimation than single image deblurring methods. On the other hand, in the image restoration step, they do not exploit the information on the noisy image, or they require *ad hoc* tuning of interdependent parameters. This paper focuses on the image restoration step and proposes a new restoration method of using a blurred/noisy image pair. In our method, the image restoration problem is formulated as a constrained convex optimization problem, where data-fidelity to a blurred image and that to a noisy image is properly taken into account as multiple hard constraints. This offers (i) high quality restoration when the blurred image also contains noise; (ii) robustness to the estimation error of the blur kernel; and (iii) easy parameter setting. We also provide an efficient algorithm for solving our optimization problem based on the so-called alternating direction method of multipliers (ADMM). Experimental results support our claims.

key words: *ADMM, deblurring, hard constraints, image restoration, constrained convex optimization*

1. Introduction

Image deblurring, removing blur from a given photograph, has been a fundamental and longstanding problem in image processing and computer vision. Many image deblurring methods, e.g., [1]–[9], are categolized as *single image blind deblurring*, that is, estimating both the blur kernel and the latent image from a single blurred image (for more information on single image blind deblurring, see a comprehensive survey [10]). Although single image blind deblurring assumes the most realistic situation, it is a very challenging task due to the highly under-cosntrained nature, so that it often leads to inaccurate estimation of blur kernels, high sensitivity to noise, and heavy dependence on prior information used.

To overcome the inherent difficulty, blind deblurring methods with a blurred/noisy image pair have been studied [11]–[15]. These methods consider such a situation that both images are captured in low light conditions with different settings. Specifically, the blurred image is taken with a slow shutter speed and a low ISO setting. With enough light, it has the correct color and intensity, but it is blurry due to camera shake. On the other hands, the noisy image is taken with a fast shutter speed and a high ISO setting. It is sharp but very noisy because of insufficient exposure and high camera gain. In addition, since it has low contrast, the colors of this image are also partially lost. Under this situation, existing methods take a two-step approach: first estimating the blur kernel from the image pair and then restoring a sharp image using the estimated kernel. Essentially, they can yield a better and much more stable result in their kernel estimation step than single image deblurring methods, since the difference between the two images is extremely informative for kernel estimation.

On the other hand, there exists a room for improvement in the image restoration step of these methods. Specifically, the methods [11]–[13] estimate the latent image based only on a given blurred image, i.e., do not exploit the information on a given noisy image in their image restoration step, so that they are sensitive to the estimation error of the blur kernel. In addition, if noise in the blurred image is not negligible, restoring a sharp image from it becomes difficult even when using the true kernel. Meanwhile, the methods [14], [15] exploit the information on both images in their image restoration step. In these methods, the image restoration problem is formulated as the minimization of a regularization term, reflecting prior information on the latent image, plus two data-fidelity terms, keeping the consistency to a blurred/noisy image pair, where the balance among these terms is controled by multiple weights. However, the tuning of such multiple weights is a tedious task because they are *interdependent* and have no physical meaning. Indeed, suitable values of them vary depending on the latent image and/or the regularization terms used.

Based on the above discussion, we propose a new image restoration method of using both a blurred image and a noisy image, which can be integrated into any blind deblurring methods with a blurred/noisy image pair. In our method, the image restoration problem is formulated as a constrained convex optimization problem: minimizing a (possibly nonsmooth) regularization function subject to multiple hard constraints. Two of the hard constraints correspond to data-fidelity to a blurred image and that to a noisy image, respectively, where the degree of fidelity to each image can be controlled by *independent* parameters that are explicitly related to the noise intensity of the image pair.

Manuscript received December 13, 2016.
Manuscript revised March 24, 2017.
Manuscript publicized June 14, 2017.
[†]The author is with the Department of Information and Communications Engineering at the Tokyo Institute of Technology, Yokohama-shi, 226–8503 Japan.
[††]The authors are with Institute of Innovative Research (IIR), Tokyo Institute of Technology, Yokohama-shi, 226–8503 Japan.
[†††]The authors are with Laboratory for Future Interdisciplinary Research of Science and Technology (FIRST), Yokohama-shi, 226–8503 Japan.
a) E-mail: takeyama.s.aa@m.titech.ac.jp
DOI: 10.1587/transinf.2016PCP0003

We also prove the existence of an optimal solution of the problem under reasonable assumptions. Since our problem formulation properly incorporates the information on a blurred/noisy image pair, it achieves (i) high quality restoration when the blurred image also contains noise; and (ii) robustness to the estimation error of the blur kernel. At the same time, the independence and clear meaning of the parameters thanks to the hard constraints offer (iii) easy parameter setting.

Through several reformulations, we also provide an efficient algorithm with guaranteed convergence for solving the constrained convex optimization problem. Our algorithm is based on the alternating direction method of multipliers (ADMM) [16]–[18], a celebrated optimization method based on proximal splitting.

The remainder of the paper is organized as follows. Section 2 introduces key tools of proximal splitting optimization used in our method. Section 3 is devoted to newly formulating a constrained convex optimization problem for image restoration using a blurred/noisy image pair and developing an ADMM-based optimization algorithm for solving it efficiently. The said three advantages of the proposed method are demonstrated through comprehensive experiments in Sect. 4. Finally, we conclude the paper in Sect. 5.

2. Preliminaries

2.1 Notations and Definitions

In this paper, let $\mathbb{R}$ be the set of real numbers. We shall use bold face lowercase and capital to represent vectors and matrices, respectively. We denote the transpose of a vector or a matrix by $(\cdot)^\top$. The standard Euclidean norm (ℓ_2 norm) of a vector is denoted by $\|\cdot\|$.

A function $f : \mathbb{R}^N \to (-\infty, \infty]$ is called *proper lower semicontinuous convex* if $\mathrm{dom}(f) := \{\mathbf{x} \in \mathbb{R}^N | \, f(\mathbf{x}) < \infty\} \neq \emptyset$, $\mathrm{lev}_{\leq\alpha}(f) := \{\mathbf{x} \in \mathbb{R}^N | \, f(\mathbf{x}) \leq \alpha\}$ is closed for every $\alpha \in \mathbb{R}$, and $f(\lambda\mathbf{x}+(1-\lambda)\mathbf{y}) \leq \lambda f(\mathbf{x})+(1-\lambda)f(\mathbf{y})$ for every $\mathbf{x}, \mathbf{y} \in \mathbb{R}^N$ and $\lambda \in (0, 1)$, respectively. Let $\Gamma_0(\mathbb{R}^N)$ be the set of all proper lower semicontinuous convex functions on $\mathbb{R}^N$.

2.2 Proximity Operator

The *proximity operator* [19] plays a central role in convex optimization based on proximal splitting. The proximity operator of $f \in \Gamma_0(\mathbb{R}^N)$ with index $\gamma > 0$ is then defined by

$$\mathrm{prox}_{\gamma f} : \mathbb{R}^N \to \mathbb{R}^N : \mathbf{x} \mapsto \operatorname*{argmin}_{\mathbf{y}} f(\mathbf{y}) + \frac{1}{2\gamma}\|\mathbf{y}-\mathbf{x}\|^2, \quad (1)$$

where the existence and uniqueness of the minimizer are guaranteed respectively by the coercivity† and the strict convexity of $f(\cdot) + \frac{1}{2\gamma}\|\cdot - \mathbf{x}\|^2$. Examples (calculations) of the proximity operator will be introduced as necessary.

We also introduce the indicator function of a nonempty closed convex set $C \subset \mathbb{R}^N$, defined by

$$\iota_C(\mathbf{x}) := \begin{cases} 0, & \text{if } \mathbf{x} \in C, \\ \infty, & \text{otherwise.} \end{cases}$$

By letting $f := \iota_C$ in (1), the proximity operator is reduced to the metric projection onto C, i.e., for any $\gamma > 0$,

$$\mathrm{prox}_{\gamma\iota_C}(\mathbf{x}) = P_C(\mathbf{x}) := \operatorname*{argmin}_{\mathbf{y}\in C} \|\mathbf{x} - \mathbf{y}\|.$$

It finds a point in C which has the minimum Euclid distance from $\mathbf{x}$.

2.3 Alternating Direction Method of Multipliers (ADMM)

Consider convex optimization problems of the form:

$$\min_{\mathbf{x},\mathbf{z}} f(\mathbf{x}) + g(\mathbf{z}) \text{ s.t. } \mathbf{z} = \mathbf{G}\mathbf{x}, \quad (2)$$

where $f \in \Gamma_0(\mathbb{R}^N)$, $g \in \Gamma_0(\mathbb{R}^M)$, and $\mathbf{G} \in \mathbb{R}^{M\times N}$. Here, we assume that f is quadratic and that g is *proximable*, i.e., the proximity operator of g is computable.

The *alternating direction method of multipliers* (ADMM) [16]–[18] is an optimization method based on proximal splitting that solves Prob. (2) by the following algorithm: for any $\mathbf{z}^{(0)}, \mathbf{d}^{(0)} \in \mathbb{R}^M$, iterate

$$\left\lfloor \begin{aligned} \mathbf{x}^{(n+1)} &= \operatorname*{argmin}_{\mathbf{x}} f(\mathbf{x}) + \tfrac{1}{2\gamma}\|\mathbf{z}^{(n)} - \mathbf{G}\mathbf{x} - \mathbf{d}^{(n)}\|^2, \\ \mathbf{z}^{(n+1)} &= \mathrm{prox}_{\gamma g}(\mathbf{G}\mathbf{x}^{(n+1)} + \mathbf{d}^{(n)}), \\ \mathbf{d}^{(n+1)} &= \mathbf{d}^{(n)} + \mathbf{G}\mathbf{x}^{(n+1)} - \mathbf{z}^{(n+1)}, \end{aligned} \right. \quad (3)$$

where $\gamma > 0$ is the step size of ADMM.

We recall the following theorem by Eckstein-Bertsekas [17], which provides a convergence property of ADMM.

Theorem 1 (Convergence of ADMM [17]). *Consider Prob.* (2)*, and assume that* $\mathbf{G}^\top\mathbf{G}$ *is invertible and that a saddle point of its unaugmented Lagrangian* $\mathcal{L}_0(\mathbf{x}, \mathbf{z}, \mathbf{y}) := f(\mathbf{x}) + g(\mathbf{z}) - \langle \mathbf{d}, \mathbf{G}\mathbf{x} - \mathbf{z}\rangle$ *exists.*†† *Then the sequence* $(\mathbf{x}_n)_{n>0}$ *generated by* (3) *converges to an optimal solution of Prob.* (2).

3. Proposed Method

3.1 Problem Formulation

Consider to estimate an unknown latent color image $\bar{\mathbf{u}} \in \mathbb{R}^{3N}$ (3 is the number of color channels, and N is the number of pixels) from an observed blurred image $\mathbf{v}_1 \in \mathbb{R}^{3N}$ and an observed noisy image $\mathbf{v}_2 \in \mathbb{R}^{3N}$. Specifically, following the prior work [14], [15] we model them as

†A function $f \in \Gamma_0(\mathbb{R}^N)$ is called *coercive* if $\|\mathbf{x}\| \to \infty \Rightarrow f(\mathbf{x}) \to \infty$. In this case, the existence of a minimizer of f is guaranteed, that is, there exists $\mathbf{x}^\star \in \mathrm{dom}(f)$ such that $f(\mathbf{x}^\star) = \inf_{\mathbf{x}\in\mathcal{H}} f(\mathbf{x})$ (see, e.g., [20]).

††A triplet $(\hat{\mathbf{x}}, \hat{\mathbf{z}}, \hat{\mathbf{d}})$ is a saddle point of an unaugmented Lagrangian $\mathcal{L}_0$ if and only if $\mathcal{L}_0(\hat{\mathbf{x}}, \hat{\mathbf{z}}, \mathbf{d}) \leq \mathcal{L}_0(\hat{\mathbf{x}}, \hat{\mathbf{z}}, \hat{\mathbf{d}}) \leq \mathcal{L}_0(\mathbf{x}, \mathbf{z}, \hat{\mathbf{d}})$, for any $(\mathbf{x}, \mathbf{z}, \mathbf{d}) \in \mathbb{R}^N \times \mathbb{R}^M \times \mathbb{R}^M$.

$$\mathbf{v}_1 = \mathbf{\Phi}\bar{\mathbf{u}} + \mathbf{n}_1, \tag{4}$$
$$\mathbf{v}_2 = \bar{\mathbf{u}} + \mathbf{n}_2, \tag{5}$$

where $\mathbf{\Phi} \in \mathbb{R}^{3N\times 3N}$ is a blur operator estimated in advance, and $\mathbf{n}_1$ and $\mathbf{n}_2$ are additive white Gaussian noises with their standard deviations σ_1 and σ_2, respectively. The model assumes that the blurred image $\mathbf{v}_1$ also contains noise (usually $\sigma_1 < \sigma_2$), which is a realistic setting as addressed in [14], [21].

Based on the above observation models, we newly formulate a convex optimization problem with multiple hard constraints for image restoration using a blurred/noisy image pair as follows:

$$\min_{\mathbf{u}} \mathcal{R}(\mathbf{\Psi}\mathbf{u}) \quad \text{s.t.} \begin{cases} \mathbf{u} \in [0,255]^{3N}, \\ \mathbf{\Phi}\mathbf{u} \in \mathcal{B}_{\mathbf{v}_1,\varepsilon_1} := \{\mathbf{x} \in \mathbb{R}^{3N} \mid \|\mathbf{x}-\mathbf{v}_1\| \leq \varepsilon_1\}, \\ \mathbf{u} \in \mathcal{B}_{\mathbf{v}_2,\varepsilon_2} := \{\mathbf{x} \in \mathbb{R}^{3N} \mid \|\mathbf{x}-\mathbf{v}_2\| \leq \varepsilon_2\}. \end{cases} \tag{6}$$

Here, $\mathcal{R}\circ\mathbf{\Psi} : \mathbb{R}^N \to (-\infty, \infty]$ is a regularization function ($\mathbf{\Psi} \in \mathbb{R}^{L\times 3N}, \mathcal{R} \in \Gamma_0(\mathbb{R}^L)$). We assume that the proximity operator of $\mathcal{R}$ (NOT $\mathcal{R}\circ\mathbf{\Psi}$) can be computed efficiently. This assumption is essential in solving the problem by ADMM, as will be explained in Sect. 3.2. The first constraint set $[0,255]^{3N} \subset \mathbb{R}^{3N}$ is the dynamic range of eight-bit color images, and the second and third ones $\mathcal{B}_{\mathbf{v}_1,\varepsilon_1}, \mathcal{B}_{\mathbf{v}_2,\varepsilon_2} \subset \mathbb{R}^{3N}$ are ℓ_2-norm balls that represent data-fidelity to a blurred image $\mathbf{v}_1$ and that to a noisy image $\mathbf{v}_2$, respectively, where $\varepsilon_1 \geq 0$ and $\varepsilon_2 \geq 0$ are their radiuses determined based on the noise intensity (noise standard deviations) of $\mathbf{v}_1$ and $\mathbf{v}_2$.

Remark 1 (Design of regularization). Total variation (TV) [22] and its vectorial variants, e.g., [23]–[27], are well-known edge-preserving regularizers for images, and they have been used in many deblurring methods. In this case, $\mathcal{R}$ is some norm, e.g., the ℓ_1 norm, the mixed $\ell_{1,2}$ norm or the nuclear norm, and $\mathbf{\Psi}$ is a discrete gradient operator. The proximity operators of such norms are computable (see, e.g., [28], [29]). Another well-known example is frame regularization relying on the sparsity of images in some transformed domain. In this case, $\mathcal{R}$ is the ℓ_1 norm, and $\mathbf{\Psi}$ is a frame analysis operator, e.g., wavelet [30] and curvelet [31]. More involved regularization, such as nonlocal regularization [32]–[34], regularization using learned operators [35], [36] and plug-and-play regularization [37], [38], can also be handled in our formulation by setting $\mathbf{\Psi}$ to the corresponding nonlocal/learned analysis operator.

Remark 2 (Benefits of incorporating data-fidelity as hard constraints). Since ε_1 and ε_2, the radiuses of the ℓ_2-norm balls in (6), are directly related to the noise intensity (noise standard deviations) of a blurred image and a noisy image, respectively, one can determine their values with the help of existing noise level estimation methods. More importantly, these parameters can be determined (almost) independent of the latent image $\bar{\mathbf{u}}$ and the regularization function $\mathcal{R}\circ\mathbf{\Psi}$. This means that once finding suitable values of ε_1 and ε_2 for some noise intensity, they can be used for various types of latent images and regularization functions under the same noise intensity, which makes the setting of parameters on data-fidelity much easier than existing methods that requires the tuning of interdependent parameters (see Sect. 4.4 for experimental validation). Such benefits of hard constraints have also been addressed, for example, in [39]–[45].

The following statement is on the existence of a solution of Prob. (6).

Proposition 1. *Assume that the intersection of the constraint sets in* (6) *is nonempty, i.e.,*

$$S := [0,255]^{3N} \cap \mathbf{\Phi}\mathcal{B}_{\mathbf{v}_1,\varepsilon_1} \cap \mathcal{B}_{\mathbf{v}_2,\varepsilon_2} \neq \emptyset,$$

and that there exists some $\mathbf{x} \in S$ *such that* $\mathcal{R}(\mathbf{\Psi}\mathbf{x}) < \infty$. *Then, Prob.* (6) *has at least one optimal solution, i.e., the function* $\mathcal{R}\circ\mathbf{\Psi}$ *has a minimizer over* S.

Proof: Since $\mathcal{R}\circ\mathbf{\Psi} \in \Gamma_0(\mathbb{R}^{3N})$ and S is a bounded closed convex subset of $\mathbb{R}^{3N}$, the statement is a direct consequence of [20, Proposition 11.14] . □

3.2 Optimization

Since Prob. (6) is a highly nonsmooth constrained problem, we need suitable iterative optimization methods to solve it. In this paper, we adopt ADMM, reviewed in Sect. 2.3. In what follows, we reformulate Prob. (6) into the ADMM-applicable form, i.e., Prob. (2).

First, let us define the indicator functions (see Sect. 2.2) of the closed convex sets $[0,255]^{3N}$, $\mathcal{B}_{\mathbf{v}_1,\varepsilon_1}$ and $\mathcal{B}_{\mathbf{v}_2,\varepsilon_2}$. Then, Prob. (6) can be rewritten as

$$\min_{\mathbf{u}} \mathcal{R}(\mathbf{\Psi}\mathbf{u}) + \iota_{[0,255]^{3N}}(\mathbf{u}) + \iota_{\mathcal{B}_{\mathbf{v}_1,\varepsilon_1}}(\mathbf{\Phi}\mathbf{u}) + \iota_{\mathcal{B}_{\mathbf{v}_2,\varepsilon_2}}(\mathbf{u}). \tag{7}$$

Second, we replace the input variables of all the terms in (7) with auxiliary variables $\mathbf{z}_1, \ldots, \mathbf{z}_4$, and express the relation between the input and the auxiliary variables by linear equality constraints, yielding

$$\min_{\mathbf{u}} \mathcal{R}(\mathbf{z}_1) + \iota_{[0,255]^{3N}}(\mathbf{z}_2) + \iota_{\mathcal{B}_{\mathbf{v}_1,\varepsilon_1}}(\mathbf{z}_3) + \iota_{\mathcal{B}_{\mathbf{v}_2,\varepsilon_2}}(\mathbf{z}_4)$$
$$\text{s.t.} \quad \mathbf{z}_1 = \mathbf{\Psi}\mathbf{u},\ \mathbf{z}_2 = \mathbf{u},\ \mathbf{z}_3 = \mathbf{\Phi}\mathbf{u},\ \mathbf{z}_4 = \mathbf{u}. \tag{8}$$

Third, we define

$$g(\mathbf{z}_1, \ldots, \mathbf{z}_4) := \mathcal{R}(\mathbf{z}_1) + \iota_{[0,255]^{3N}}(\mathbf{z}_2) + \iota_{\mathcal{B}_{\mathbf{v}_1,\varepsilon_1}}(\mathbf{z}_3) + \iota_{\mathcal{B}_{\mathbf{v}_2,\varepsilon_2}}(\mathbf{z}_4).$$

Then, the function g becomes proximable thanks to the variable splitting, as long as each term of g is proximable. Indeed, $\mathcal{R}$ is proximable from the assumption, and the other terms, the three indicator functions, are also proximable because the metric projections onto the corresponding closed convex sets are available (see Sect. 2.2, (10) and (11)).

Finally, by letting $\mathbf{I}$ be the identity matrix of size $3N \times 3N$ and defining

$$f(\mathbf{u}) := 0 \quad \text{and} \quad \mathbf{G} := \begin{pmatrix} \mathbf{\Psi} \\ \mathbf{I} \\ \mathbf{\Phi} \\ \mathbf{I} \end{pmatrix}, \tag{9}$$

Fig. 1 Test images

Prob. (8) is reduced to Prob. (2).

The resulting algorithm based on ADMM is summarized in Alg. 1. Since $\mathbf{G}$ in (9) is a full column rank matrix due to $\mathbf{I}$, $\mathbf{G}^\top\mathbf{G}$ is invertible, so that the convergence of Alg. 1 is guaranteed if a saddle point of $g(\mathbf{z}) - \langle \mathbf{d}, \mathbf{G}\mathbf{u} - \mathbf{z}\rangle$ exists.

Now we discuss the computation of each step of Alg 1. Since the update of $\mathbf{u}$ (Step 2) is strictly-convex quadratic minimization because of the full-column-rankness of $\mathbf{G}$, it boils down to solving the following matrix inversion:

$$\begin{aligned}\mathbf{u}^{(n+1)} =& (\mathbf{\Psi}^\top\mathbf{\Psi} + \mathbf{\Phi}^\top\mathbf{\Phi} + 2\mathbf{I})^{-1}\mathrm{RHS}\\ \mathrm{RHS} :=& (\mathbf{\Psi}^\top(\mathbf{z}_1^{(n)} - \mathbf{d}_1^{(n)}) + (\mathbf{z}_2^{(n)} - \mathbf{d}_2^{(n)})\\ &+ \mathbf{\Phi}^\top(\mathbf{z}_3^{(n)} - \mathbf{d}_3^{(n)}) + (\mathbf{z}_4^{(n)} - \mathbf{d}_4^{(n)})).\end{aligned}$$

If the matrix $(\mathbf{\Psi}^\top\mathbf{\Psi} + \mathbf{\Phi}^\top\mathbf{\Phi} + 2\mathbf{I})$ is a block circulant with circulant blocks (BCCB) matrix [46], we can leverage 2D fast Fourier transform to efficiently solve the inversion (in $O(N \log N)$ time) because the matrix can be diagonalized by the 2D dicrete Fourier transform matrix. The matrix $\mathbf{\Phi}^\top\mathbf{\Phi}$ becomes a BCCB matrix provided that the corresponding blur kernel is spatially invariant. Meanwhile, the structure of $\mathbf{\Psi}^\top\mathbf{\Psi}$ depends on the design of regularization. If $\mathbf{\Psi}$ is a discrete gradient operator with periodic boundary (e.g., TV regularization), $\mathbf{\Psi}^\top\mathbf{\Psi}$ becomes a BCCB matrix. If $\mathbf{\Psi}$ is a Parseval tight frame [47] (e.g., wavelet/curvelet regularization), $\mathbf{\Psi}^\top\mathbf{\Psi} = \mathbf{I}$. Thus for these cases, the 2DFFT-based computation is possible. Otherwise, we offer to use a preconditioned conjugate gradient method [48] for approximately solving the inversion.

For the updates of $\mathbf{z}_1, \dots, \mathbf{z}_4$ (Step 3-6), we need to compute the proximity operators of each term of g. The proximity operator of $\mathcal{R}$ depends on the design of regularization. As addressed in Remark 1, it is indeed computable for many types of regularization. The proximity operator of $\iota_{[0,255]^{3N}}$ equals to the metric projection onto the box constraint $[0, 255]^{3N}$, given, for $i = 1, \dots, 3N$, by

$$[P_{[0,255]^{3N}}(\mathbf{x})]_i = \begin{cases} 0, & \text{if } x_i < 0,\\ 255, & \text{if } x_i > 255,\\ x_i & \text{otherwise.}\end{cases} \tag{10}$$

The proximity operators of $\iota_{\mathcal{B}_{\mathbf{v}_1,\varepsilon_1}}$ and $\iota_{\mathcal{B}_{\mathbf{v}_2,\varepsilon_2}}$ can also be computed by the metric projection onto a $\mathbf{v}$-centered ℓ_2-norm ball with radius $\varepsilon \geq 0$, given by

---

Algorithm 1: ADMM method for Prob. (6)

input : $\mathbf{z}_1^{(0)}, \mathbf{z}_2^{(0)}, \mathbf{z}_3^{(0)}, \mathbf{z}_4^{(0)}, \mathbf{d}_1^{(0)}, \mathbf{d}_2^{(0)}, \mathbf{d}_3^{(0)}, \mathbf{d}_4^{(0)}$, and $\gamma > 0$

1 **while** *A stopping criterion is not satisfied* **do**

2 $\quad \mathbf{u}^{(n+1)} = \operatorname*{argmin}_{\mathbf{u}} \frac{1}{2\gamma}(\|\mathbf{z}_1^{(n)} - \mathbf{\Psi}\mathbf{u}^{(n)} - \mathbf{d}_1^{(n)}\|^2 + \|\mathbf{z}_2^{(n)} - \mathbf{u}^{(n)} - \mathbf{d}_2^{(n)}\|^2 + \|\mathbf{z}_3^{(n)} - \mathbf{\Phi}\mathbf{u}^{(n)} - \mathbf{d}_3^{(n)}\|^2 + \|\mathbf{z}_4^{(n)} - \mathbf{u}^{(n)} - \mathbf{d}_4^{(n)}\|^2)$;

3 $\quad \mathbf{z}_1^{(n+1)} = \mathrm{prox}_{\gamma\mathcal{R}}(\mathbf{\Psi}\mathbf{u}^{(n+1)} + \mathbf{d}_1^{(n)})$;

4 $\quad \mathbf{z}_2^{(n+1)} = \mathrm{prox}_{\gamma\iota_{[0,255]^{3N}}}(\mathbf{u}^{(n+1)} + \mathbf{d}_2^{(n)})$;

5 $\quad \mathbf{z}_3^{(n+1)} = \mathrm{prox}_{\gamma\iota_{\mathcal{B}_{\mathbf{v}_1,\varepsilon_1}}}(\mathbf{\Phi}\mathbf{u}^{(n+1)} + \mathbf{d}_3^{(n)})$;

6 $\quad \mathbf{z}_4^{(n+1)} = \mathrm{prox}_{\gamma\iota_{\mathcal{B}_{\mathbf{v}_2,\varepsilon_2}}}(\mathbf{u}^{(n+1)} + \mathbf{d}_4^{(n)})$;

7 $\quad \mathbf{d}_1^{(n+1)} = \mathbf{d}_1^{(n)} + \mathbf{\Psi}\mathbf{u}^{(n+1)} - \mathbf{z}_1^{(n+1)}$;

8 $\quad \mathbf{d}_2^{(n+1)} = \mathbf{d}_2^{(n)} + \mathbf{u}^{(n+1)} - \mathbf{z}_2^{(n+1)}$;

9 $\quad \mathbf{d}_3^{(n+1)} = \mathbf{d}_3^{(n)} + \mathbf{\Phi}\mathbf{u}^{(n+1)} - \mathbf{z}_3^{(n+1)}$;

10 $\quad \mathbf{d}_4^{(n+1)} = \mathbf{d}_4^{(n)} + \mathbf{u}^{(n+1)} - \mathbf{z}_4^{(n+1)}$;

11 $\quad n \leftarrow n + 1$;

$$P_{\mathcal{B}_{\mathbf{v},\varepsilon}}(\mathbf{x}) = \begin{cases} \mathbf{x}, & \text{if } \mathbf{x} \in \mathcal{B}_{\mathbf{v},\varepsilon},\\ \mathbf{v} + \frac{\varepsilon(\mathbf{x}-\mathbf{v})}{\|\mathbf{x}-\mathbf{v}\|}, & \text{otherwise.}\end{cases} \tag{11}$$

4. Experiments

4.1 Experimental Setting

We demonstrate the three advantages of the proposed method (see Sect. 1) through three experiments. In the following experiments, we used 20 color images used as test images, which are taken from the *Berkley Segmentation Database* [49] (Fig. 1).

We utilized a popular color TV [23] as regularization in our method, which is defined as follows:

$$\mathrm{TV}(\mathbf{u}) := \|\mathbf{D}\mathbf{u}\|_{1,2},$$

where $\mathbf{D}$ is a discreat gradient operator defined as

$$\mathbf{D} : \mathbb{R}^{3N} \to \mathbb{R}^{6N} : \mathbf{u} \mapsto (\mathbf{d}_v^\top \mathbf{d}_h^\top)^\top$$

with $\mathbf{d}_v, \mathbf{d}_h \in \mathbb{R}^{3N}$ being the vertical and horizontal differences of a color image $\mathbf{u}$, and $\|\cdot\|_{1,2}$ is the mixed $\ell_{1,2}$ norm defined as

$$\|\cdot\|_{1,2} : \mathbb{R}^{6N} \to \mathbb{R} : \mathbf{x} \mapsto \sum_{i=1}^{N} \sqrt{\sum_{j=0}^{5} x_{i+jN}^2}.$$

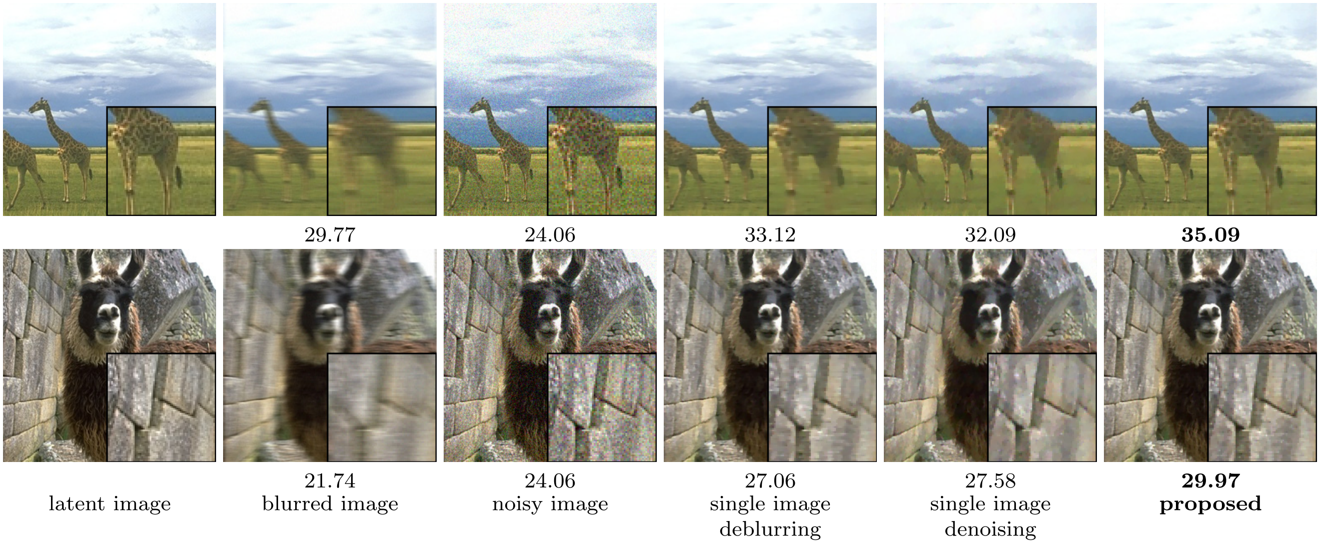

Fig. 3 Resulting images with their PSNR in the first experiment ($\sigma_2/\sigma_1 = 8$).

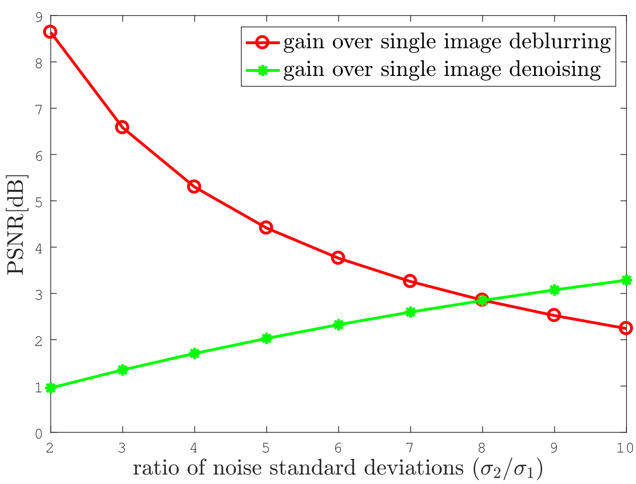

Fig. 2 PSNR gain of the proposed method over the single image deblurring/denoising.

In this case, $\mathcal{R} := \|\cdot\|_{1,2}$ and $\mathbf{\Psi} := \mathbf{D}$ in Prob. (6). The proximity operator of the mixed $\ell_{1,2}$ norm can be computed by a simple soft-thresholding type operation: for $\gamma > 0$ and for $i = 1, \ldots, 6N$,

$$[\mathrm{prox}_{\gamma\|\cdot\|_{1,2}}(\mathbf{x})]_i := \max\left\{1 - \gamma\left(\sum_{j=0}^{5} x_{\tilde{i}+jN}^2\right)^{-\frac{1}{2}}, 0\right\} x_i,$$

where $\tilde{i} := ((i-1) \mod N) + 1$.

We set ε_1 and ε_2 in (6) to $0.95\sqrt{3N\sigma_1^2}$ and $0.95\sqrt{3N\sigma_2^2}$, respectively.

We adopted the peak signal-to-noise ratio (PSNR) [dB] to evaluate the objective quality of a restored image $\mathbf{u}$, which is given by

$$20 \cdot \log 10 \frac{3N \times 255}{\|\mathbf{u} - \bar{\mathbf{u}}\|}.$$

4.2 Basic Performance Evaluation

To evaluate the effectiveness of incorporating multiple hard constraints on data-fidelity to a blurred/noisy image pair, we compare our method with two single image restoration methods. One is single image deblurring, i.e., solving Prob. (6) without the constraint $\mathbf{u} \in \mathcal{B}_{\mathbf{v}_2,\varepsilon_2}$. The other is single image denoising, i.e., solving Prob. (6) without the constraint $\mathbf{\Phi u} \in \mathcal{B}_{\mathbf{v}_1,\varepsilon_1}$. As in our method, ADMM was used for both methods.

In this experiment, we generated blurred images as follows: clean test images are blurred by a horizontal motion blur of 9 pixels and then contaminated by an additive white Gaussian noise $\mathbf{n}_1$ in (4) with the standard deviation $\sigma_1 = 2$. Meanwhile, noisy images were generated by adding a white Gaussian noise $\mathbf{n}_2$ in (5) to clean test images, where the standard deviation σ_2 was increased from 4 to 20 by 2.

Figure 2 plots the PSNR gain of our method over the single image deblurring (circle marker) and the single image denoising (asterisk marker), where PSNR is averaged over the 20 test images. One can see that for all the ratio of the noise standard deviations σ_2/σ_1, the proposed method outperforms both single image deblurring and denoising methods. This observation suggests that for a given blurred/noisy image pair, exploiting information on the noisy image in the image restoration step is very effective when the blurred image also contains noise.

Figure 3 depicts some resulting images with their PSNR ($\sigma_2/\sigma_1 = 8$). One can see that 1. details are lost in the images obtained by the single image deblurring, 2. color artifact remains in the images obtained by the single image denoising, and 3. Our method achieves detail-preserving restoration with much less artifact.

We also check the convergence behavior of our algorithm (Alg. 1). For evaluation of convergence, we define the normalized root mean square error (NRMSE) between the

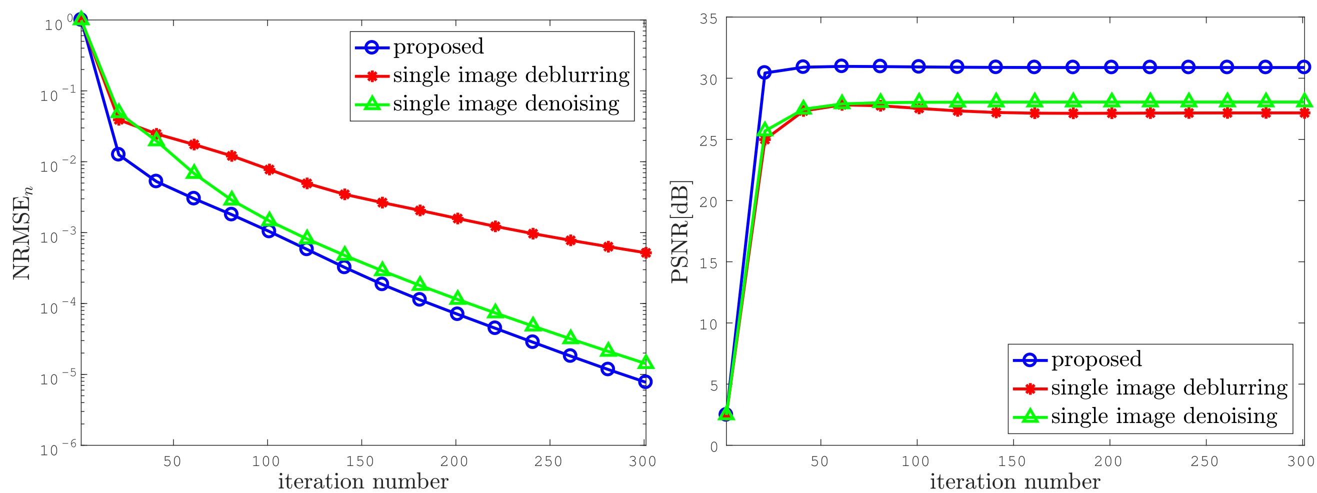

Fig. 4 Evolution of NRMSE$_n$ versus iterations (left) and the evolution of PSNR[dB] versus iterations (right) of Alg. 1.

Table 1 Results of the experiment using inaccurate blur kernel.

kernel error	0°	1°	2°	3°	4°	5°	6°	7°	8°	9°	10°	spatially-varying
PSNR [dB]	32.97	32.91	32.62	32.29	32.01	31.79	31.61	31.47	31.36	31.26	31.17	30.12
gain over single image deblurring	2.85	2.88	3.06	3.40	3.81	4.18	4.47	4.70	4.88	5.03	5.15	6.58

current estimate $\mathbf{u}^{(n)}$ and the optimal solution $\mathbf{u}^\star$ of Prob. (6), i.e., NRMSE$_n$:= $\|\mathbf{u}^{(n)} - \mathbf{u}^\star\|/\|\mathbf{u}^\star\|$. Since the optimal solution $\mathbf{u}^\star$ is analytically unavailable, it was pre-computed by Alg. 1 with 100000 iterations. Figure 4 plots the evolution of NRMSE$_n$ versus iterations (left) and the evolution of PSNR versus iterations (right), where the stepsize γ of ADMM was set to 0.01. These plots suggest that Alg. 1 properly works, and that exploiting the information on a blurred/noisy image pair makes the convergence of ADMM faster than the single image deblurring/denoising, which is a positive side effect of using a blurred/noisy image pair.

4.3 Robustness to Inaccurate Blur Kernels

To illustrate the robustness of our method to the estimation error of blur kernels, we conducted the following experiment. First, we generated images blurred by a certain blur kernel, which we refer to as the true blur kernel, and then in the image restoration step, we used an inaccurate blur kernel. Specifically, we consider the two cases: motion blur and spatially-varying blur. In the motion blur case, the true blur kernel was set to a horizontal motion blur of 9 pixels, and the inaccurate blur kernel was set to a motion blur of 9 pixels with its angle $\theta > 0$, where we examined $\theta = 1°$ to $10°$ by $1°$. In the spatially-varying blur case, the true blur matrix $\mathbf{\Phi}$ was made from spatially-varying per-pixel kernels, as visualized in Fig. 6, and the inaccurate blur kernel was set to be spatially invariant with its kernel being the center kernel of the second image from left in Fig. 6. For both cases, the blurred images contain an additive white Gaussian noise with the standard deviation $\sigma_1 = 2$, and the noisy images $\sigma_2 = 16$.

Table 1 shows PSNR of restored images and the PSNR gain over the single image deblurring, where these valuers are averaged over the 20 test images. One can observe that for the motion blur case, the PSNR gain over the single image deblurring is proportional to the angle error, implying the robustness of our method to inaccurate blur kernels compared with the single image deblurring. For the spartially-varying blur case, the PSNR gain is also significant.

Figure 5 and Fig. 6 depict several resulting images with their PSNR. One can see that the single image deblurring leads to oversmoothing when the kernel error is large. By contrast, our method can restore sharp images in such a situation.

4.4 Facilitation of Parameter Setting

In the final experiment, we demonstrate that the setting of the parameters on data-fidelity in our method are much easier than the existing methods using a blurred/noisy image pair in the restoration step [14], [15]. In the existing methods, image restoration is performed by minimizing the sum of some specific regularization term and two data-fidelity terms on a blurred/noisy image pair, which can be expressed as the following optimization problem:

$$\min_{\mathbf{u}} \mathcal{R}(\mathbf{\Psi u}) + \frac{\lambda_1}{2}\|\mathbf{\Phi u} - \mathbf{v}_1\|_2^2 + \frac{\lambda_2}{2}\|\mathbf{u} - \mathbf{v}_2\|_2^2, \qquad (12)$$

where $\lambda_1, \lambda_2 > 0$ control the balance among the three terms.

Table 2 shows the (hand-optimized) best values of λ_1 and λ_2 (in terms of PSNR) in Prob. (12) for each test image, where the regularization term was set to the color TV. For every test image, the noise standard deviations of a

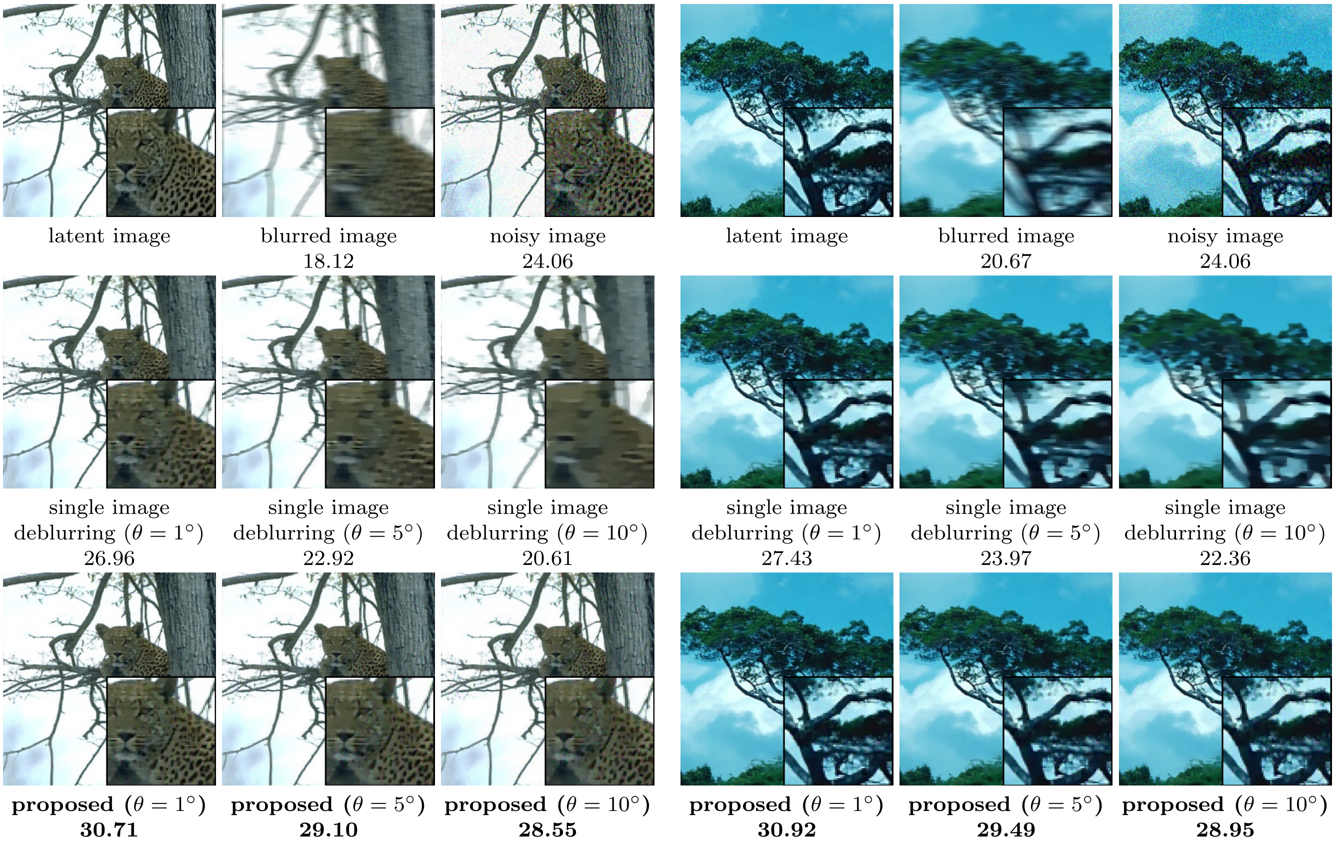

Fig. 5 Restored results with their PSNR by using inaccurate blur kernels in the motion blur case.

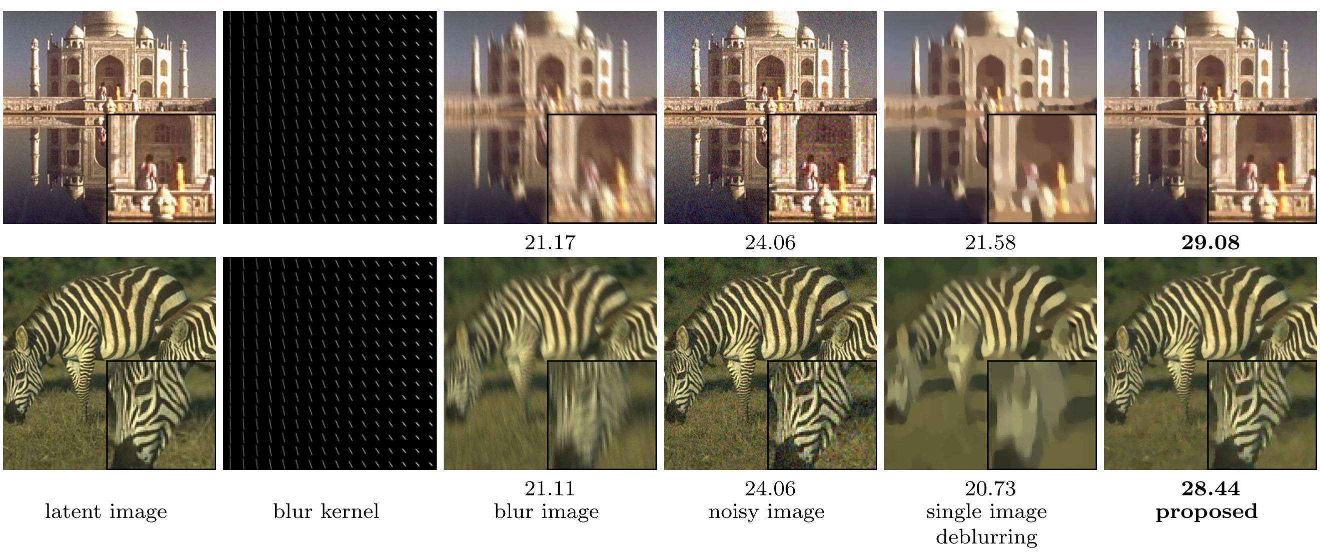

Fig. 6 Restored results with their PSNR by using an inaccurate blur kernel in the spatially-varying blur case.

Table 2 Best values of λ_1 and λ_2 in the existing formulation (Prob. (12)).

image	img1	img2	img3	img4	img5	img6	img7	img8	img9	img10
λ_1	733	1047	578	603	563	801	916	446	589	645
λ_2	8	9	11	7	9	8	7	9	6	7
image	img11	img12	img13	img14	img15	img16	img17	img18	img19	img20
λ_1	770	895	988	932	905	812	917	678	920	650
λ_2	6	8	8	8	7	6	8	8	8	8

blurred/noisy image pair (σ_1, σ_2) were fixed at $(2, 16)$. One can see that the best values of λ_1 and λ_2 are different for each test image even though the noise standard deviations are the same for every test image. This is because λ_1 and λ_2 depend on the latent image, so that the tuning of them is very difficult. By contrast, the proposed method achieves almost the same restoration performance with common parameters $\varepsilon_1 = 0.95\sqrt{3N\sigma_1^2}$ and $\varepsilon_2 = 0.95\sqrt{3N\sigma_2^2}$ for all the test images, i.e., the parameter setting is much easier.

5. Conclusion

We have proposed a new image restoration method that fully exploits the information on a blurred/noisy image pair. We formulated the image restoration problem as a convex optimization problem with multiple hard constraints, where data-fidelity measures to both a blurred image and a noisy image are imposed via ℓ_2-norm balls. Then we developed an ADMM-based algorithm for solving the problem efficiently. Our method has the three advantages over existing methods with a blurred/noisy image pair, that is, (i) high quality restoration when the blurred image also contains noise; (ii) robustness to the estimation error of the blur kernel; and (iii) easy parameter setting. We illustrated the effectiveness and utility of our method through comprehensive experiments.

Finally, we remark again that our method can be plugged into the image restoration step of any blind deblurring methods with a blurred/noisy image pair, which would enhance them. Incorporating variational image decomposition models, e.g., [50]–[53], into the proposed method is an interesting direction of future work. Also, a stochastic image restoration methodology [54] would be able to further accelerate our method.

References

[1] Q. Shan, J. Jia, and A. Agarwala, “High-quality motion deblurring from a single image,” ACM Trans. Graph., vol.27, no.3, pp.73:1–73:10, 2008.

[2] S. Cho and S. Lee, “Fast motion deblurring,” ACM Trans. Graph., vol.28, no.5, pp.145:1–145:8, 2009.

[3] J.-F. Cai, H. Ji, C. Liu, and Z. Shen, “Blind motion deblurring from a single image using sparse approximation,” Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), pp.104–111, 2009.

[4] L. Xu and J. Jia, “Two-phase kernel estimation for robust motion deblurring,” Proc. Eur. Conf. Comput. Vis. (ECCV), vol.6311, pp.157–170, 2010.

[5] D. Krishnan, T. Tay, and R. Fergus, “Blind deconvolution using a normalized sparsity measure,” Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), pp.233–240, 2011.

[6] J.-F. Cai, H. Ji, C. Liu, and Z. Shen, “Framelet-based blind motion deblurring from a single image,” IEEE Trans. Image Process., vol.21, no.2, pp.562–572, 2012.

[7] L. Zhong, S. Cho, D. Metaxas, S. Paris, and J. Wang, “Handling noise in single image deblurring using directional filters,” Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), pp.612–619, 2013.

[8] T. Michaeli and M. Irani, “Blind deblurring using internal patch recurrence,” Proc. Eur. Conf. Comput. Vis. (ECCV), vol.8691, pp.783–798, 2014.

[9] D. Perrone and P. Favaro, “A clearer picture of total variation blind deconvolution,” IEEE Trans. Pattern Anal. Mach. Intell., vol.38, no.6, pp.1041–1055, 2016.

[10] W.-S. Lai, J.-B. Huang, Z. Hu, N. Ahuja, and M.-H. Yang, “A comparative study for single image blind deblurring,” Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), pp.1701–1709, 2016.

[11] L. Yuan, J. Sun, L. Quan, and H.-Y. Shum, “Image deblurring with blurred/noisy image pairs,” ACM Trans. Graph., vol.26, no.3, pp.1:1–1:10, 2007.

[12] S.-H. Lee, H.-M. Park, and S.-Y. Hwang, “Motion deblurring using edge map with blurred/noisy image pairs,” Optics Communications, vol.285, no.7, pp.1777–1786, 2012.

[13] C. Je, H.S. Jeon, C.-H. Son, and H.-M. Park, “Disparity-based space-variant image deblurring,” Signal Process.: Image Commun., vol.28, no.7, pp.792–808, 2013.

[14] H. Li, Y. Zhang, J. Sun, and D. Gong, “Joint motion deblurring with blurred/noisy image pair,” Proc. Int. Conf. Pattern Recognit. (ICPR), pp.1020–1024, 2014.

[15] C.-H. Son and X.-P. Zhang, “Layer-based approach for image pair fusion,” IEEE Trans. Image Process., vol.25, no.6, pp.2866–2881, 2016.

[16] D. Gabay and B. Mercier, “A dual algorithm for the solution of nonlinear variational problems via finite elements approximations,” Comput. Math. Appl., vol.2, pp.17–40, 1976.

[17] J. Eckstein and D.P. Bertsekas, “On the Douglas—Rachford splitting method and the proximal point algorithm for maximal monotone operators,” Math. Program., vol.55, no.1-3, pp.293–318, 1992.

[18] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, “Distributed optimization and statistical learning via the alternating direction method of multipliers,” Foundations and Trends in Machine Learning, vol.3, no.1, pp.1–122, 2011.

[19] J.J. Moreau, “Fonctions convexes duales et points proximaux dans un espace hilbertien,” C. R. Acad. Sci. Paris Ser. A Math., vol.255, pp.2897–2899, 1962.

[20] H.H. Bauschke and P.L. Combettes, Convex analysis and monotone operator theory in Hilbert spaces, Springer, New York, 2011.

[21] L. Zhong, S. Cho, D. Metaxas, S. Paris, and J. Wang, “Handling noise in single image deblurring using directional filters,” Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), pp.612–619, 2013.

[22] L.I. Rudin, S. Osher, and E. Fatemi, “Nonlinear total variation based noise removal algorithms,” Phys. D, vol.60, no.1-4, pp.259–268, 1992.

[23] X. Bresson and T.F. Chan, “Fast dual minimization of the vectorial total variation norm and applications to color image processing,” Inverse Probl. Imag., vol.2, no.4, pp.455–484, 2008.

[24] B. Goldluecke, E. Strekalovskiy, and D. Cremers, “The natural vectorial total variation which arises from geometric measure theory,” SIAM J. Imag. Sci., vol.5, no.2, pp.537–563, 2012.

[25] S. Ono and I. Yamada, “Decorrelated vectorial total variation,” Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), pp.4090–4097, 2014.

[26] S. Lefkimmiatis, A. Roussos, P. Maragos, and M. Unser, “Structure tensor total variation,” SIAM J. Imag. Sci., vol.8, no.2, pp.1090–1122, 2015.

[27] S. Ono, K. Shirai, and M. Okuda, “Vectorial total variation based on arranged structure tensor for multichannel image restoration,” Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP), pp.4528–4532, 2016.

[28] P.L. Combettes and J.-C. Pesquet, “Proximal splitting methods in signal processing,” Fixed-Point Algorithms for Inverse Problems in Science and Engineering, vol.49, pp.185–212, Springer-Verlag, 2011.

[29] S. Ono, “Distributed convex optimization via proximal splitting: A survey on admm-based approaches,” Journal of The Society of Instrument and Control Engineers, vol.55, no.11, pp.954–959, 2016. (in Japanese).

[30] S. Mallat, A wavelet tour of signal processing, 2nd ed., Academic Press, 1999.
[31] E. Candès, L. Demanet, D. Donoho, and L. Ying, "Fast discrete curvelet transforms," SIAM J. Multi. Model. Simul., vol.5, no.3, pp.861–899, 2006.
[32] G. Gilboa and S. Osher, "Nonlocal linear image regularization and supervised segmentation," Multiscale Model. Simul., vol.6, no.2, pp.595–630, 2007.
[33] A. Danielyan, V. Katkovnik, and K. Egiazarian, "BM3D frames and variational image deblurring," IEEE Trans. Image Process., vol.21, no.4, pp.1715–1728, 2012.
[34] G. Chierchia, N. Pustelnik, B. Pesquet-Popescu, and J.-C. Pesquet, "A nonlocal structure tensor-based approach for multicomponent image recovery problems," IEEE Trans. Image Process., vol.23, no.12, pp.5531–5544, 2014.
[35] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman, "Non-local sparse models for image restoration," Proc. IEEE Int. Conf. Comput. Vis. (ICCV), pp.2272–2279, 2009.
[36] D. Zoran and Y. Weiss, "From learning models of natural image patches to whole image restoration," Proc. IEEE Int. Conf. Comput. Vis. (ICCV), pp.479–486, 2011.
[37] S. Sreehari, S.V. Venkatakrishnan, B. Wohlberg, G.T. Buzzard, L.F. Drummy, J.P. Simmons, and C.A. Bouman, "Plug-and-play priors for bright field electron tomography and sparse interpolation," IEEE Trans. Comput. Imag., vol.2, no.4, pp.408–423, 2016.
[38] S. Ono, "Primal-dual plug-and-play image restoration," IEEE Signal Process. Lett., vol.24, no.8, pp.1108–1112, 2017.
[39] M.V. Afonso, J.M. Bioucas-Dias, and M.A.T. Figueiredo, "An augmented lagrangian approach to the constrained optimization formulation of imaging inverse problems," IEEE Trans. Image Process., vol.20, no.3, pp.681–695, 2011.
[40] M. Carlavan and L. Blanc-Feraud, "Sparse Poisson noisy image deblurring," IEEE Trans. Image Process., vol.21, no.4, pp.1834–1846, 2012.
[41] T. Teuber, G. Steidl, and R.H. Chan, "Minimization and parameter estimation for seminorm regularization models with I-divergence constraints," Inverse Problems, vol.29, no.3, p.035007, 2013.
[42] G. Chierchia, N. Pustelnik, J.-C. Pesquet, and B. Pesquet-Popescu, "Epigraphical projection and proximal tools for solving constrained convex optimization problems," Signal, Image and Video Process., vol.9, no.8, pp.1737–1749, 2015.
[43] S. Ono and I. Yamada, "Signal recovery with certain involved convex data-fidelity constraints," IEEE Trans. Signal Process., vol.63, no.22, pp.6149–6163, 2015.
[44] S. Ono, "L_0 gradient projection," IEEE Trans. Image Process., vol.26, no.4, pp.1554–1564, 2017.
[45] S. Takeyama, S. Ono, and I. Kumazawa, "Hyperspectral image restoration by hybrid spatio-spectral total variation," Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP), pp.4586–4590, 2017.
[46] P.C. Hansen, J.G. Nagy, and D.P. O'Leary, Deblurring Images: Matrices, Spectra, and Filtering, SIAM, 2006.
[47] J. Kovacevic and A. Chebira, "Life beyond bases: The advent of frames (Part I)," IEEE Signal Process. Magazine, vol.24, no.4, pp.86–104, 2007.
[48] G.H. Golub and C.F.V. Loan, Matrix Computations, 4th ed., Johns Hopkins University Press, 2012.
[49] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," Proc. IEEE Int. Conf. Comput. Vis. (ICCV), pp.416–423, 2001.
[50] J.-F. Aujol, G. Gilboa, T. Chan, and S. Osher, "Structure-texture image decomposition—modeling, algorithms, and parameter selection," Int. J. Comput. Vis., vol.67, no.1, pp.111–136, 2006.
[51] V. Duval, J.-F. Aujol, and L.A. Vese, "Mathematical modeling of textures: application to color image decomposition with a projected gradient algorithm," J. Math. Imag. Vis., vol.37, no.3, pp.232–248, 2010.
[52] S. Ono, T. Miyata, I. Yamada, and K. Yamaoka, "Image recovery by decomposition with component-wise regularization," IEICE Trans. Fundamentals., vol.E95-A, no.12, pp.2470–2478, 2012.
[53] S. Ono, T. Miyata, and I. Yamada, "Cartoon-texture image decomposition using blockwise low-rank texture characterization," IEEE Trans. Image Process., vol.23, no.3, pp.1128–1142, 2014.
[54] S. Ono, M. Yamagishi, T. Miyata, and I. Kumazawa, "Image restoration using a stochastic variant of the alternating direction method of multipliers," Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP), pp.4523–4527, 2016.

Saori Takeyama received the B.E. degree in engineering in 2016 from the Tokyo Institute of Technology, Tokyo, Japan. Since 2016, she has been a master course student in the Department of Information and Communications Engineering at the Tokyo Institute of Technology. She received IE Award in 2016 from IEICE Technical Committee on Image Engineering. Her current research interests are in hyperspectral image restoration and mathematical optimization.

Shunsuke Ono received the B.E. degree in computer science in 2010, and the M.E. and Ph.D. degrees in communications and computer engineering in 2012 and 2014 from the Tokyo Institute of Technology, respectively. He is currently an Assistant Professor with Laboratory for Future Interdisciplinary Research of Science and Technology (FIRST), Institute of Innovative Research (IIR), Tokyo Institute of Technology. Since 2016, he has also been a Researcher with Precursory Research for Embryonic Science and Technology (PRESTO), Japan Science and Technology Corporation (JST), Tokyo, Japan. His research interests are in image processing and low-level computer vision, signal processing, and mathematical optimization. From April 2012 to September 2014, he was a recipient of the Research Fellowship of the Japan Society for the Promotion of Science (JSPS). He received the Excellent Paper Award in 2014 and the Young Researchers' Award in 2013 from the IEICE; the Outstanding Student Journal Paper Award from IEEE Signal Processing Society (SPS) Japan Chapter in 2014; the Yasujiro Niwa Outstanding Paper Award from Tokyo Denki University in 2015; the TELECOM System Technology Award from the Telecommunications Advancement Foundation in 2016; and the Funai Research Award from the Funai Foundation in 2017.

Itsuo Kumazawa received the Dr. Degree in Engineering from Tokyo Institute of Technology, Japan, in 1986. He joined Tokyo Institute of Technology as Research Associate in 1986 and is now Professor of Laboratory for Future Interdisciplinary Research of Science and Technology (FIRST), Institute of Innovative Research (IIR), Tokyo Institute of Technology. Dr. Kumazawa has published a number of papers and books concerning user interface, image processing, neural computing and pattern recognition. He is a member of International Neural Network Society and Institute for Electronics, Information and Communication Engineers in Japan.