Multi-Image Blind Deblurring Using a Coupled Adaptive Sparse Prior

Haichao Zhang†, David Wipf ‡ and Yanning Zhang†
† School of Computer Science, Northwestern Polytechnical University, Xi'an, China
‡ Microsoft Research Asia, Beijing, China
{hczhang1,davidwipf}@gmail.com ynzhang@nwpu.edu.cn

Abstract

This paper presents a robust algorithm for estimating a single latent sharp image given multiple blurry and/or noisy observations. The underlying multi-image blind deconvolution problem is solved by linking all of the observations together via a Bayesian-inspired penalty function which couples the unknown latent image, blur kernels, and noise levels together in a unique way. This coupled penalty function enjoys a number of desirable properties, including a mechanism whereby the relative-concavity or shape is adapted as a function of the intrinsic quality of each blurry observation. In this way, higher quality observations may automatically contribute more to the final estimate than heavily degraded ones. The resulting algorithm, which requires no essential tuning parameters, can recover a high quality image from a set of observations containing potentially both blurry and noisy examples, without knowing a priori *the degradation type of each observation. Experimental results on both synthetic and real-world test images clearly demonstrate the efficacy of the proposed method.*

1. Introduction

In many practical scenarios we are presented with multiple captures of the same physical scene under different imaging conditions. For example, this is possible using the exposure bracketing or burst-mode functionality on many consumer cameras. However, each observed image may often contain unknown, corrupting artifacts such as blurring and/or noise. A typical factor causing blur is the relative motion between camera and scene during the exposure period, which may arise from hand jitter [6, 14]. Multi-image blind deconvolution algorithms are designed to jointly utilize all available observations to produce a single sharp estimate of the underlying scene.

Given L corrupted versions of a latent sharp image $\mathbf{x}$, the uniform convolutional blur model [6] assumes the observation process

$$\mathbf{y}_l = \mathbf{k}_l * \mathbf{x} + \mathbf{n}_l, \qquad \forall l \in \{1, \ldots, L\}, \tag{1}$$

where $\mathbf{k}_l$ is a Point Spread Function (PSF) or blur kernel, $*$ denotes the convolution operator, and $\mathbf{n}_l$ is a zero-mean Gaussian noise term with covariance $\lambda_l \mathbf{I}$. Within this context, the ultimate goal of multi-image blind deblurring is to estimate the sharp and clean image $\mathbf{x}$ given only the blurry and noisy observations $\{\mathbf{y}_l\}_{l=1}^{L}$, without any prior knowledge regarding the unknown kernels $\mathbf{k}_l$ or noise levels λ_l (see Figure 1). By combining the complementary information from multiple images, it is often possible to generate higher quality estimates of the scene $\mathbf{x}$ than in the single-image, blind deconvolution case [12].

Figure 1. Dual motion deblurring examples. Full images are shown in Figure 6.

While a number of successful multi-image blind deconvolution methods exist, e.g., [12, 4, 3, 13, 22], there remains room for practical improvements and additional theoretical understanding. In this context, we present a principled energy-minimization algorithm that can handle a flexible number of degraded observations without requiring that we know the nature (e.g., blurry vs. noisy) or extent of the degradation for each observation. The underlying cost function relies on a coupled penalty function, which combines the latent sharp image estimate with a separate blur kernel and noise variance associated with each observed image. Theoretical analysis reveals that this penalty provides a useful agency for balancing the effects of observations with varying quality, while at the same time avoiding suboptimal local minima. All unknown quantities are optimized using a majorization-minimization algorithm that requires no tuning parameters. Additionally, when only a single observa-

1063-6919/13 $26.00 © 2013 IEEE
DOI 10.1109/CVPR.2013.140

tion is present, the method reduces to a principled, single-image blind deconvolution algorithm with an image penalty that adaptively interpolates between the ℓ_0 and ℓ_1 norms. Experimental results on both synthetic and real-world test images validate the proposed method relative to the current state-of-the-art.

The remainder of the paper is organized as follows. Section 2 briefly reviews existing multi-image blind deconvolution algorithms; we then introduce our alternative algorithm in Section 3. Theoretical properties and analysis related to the proposed coupled penalty function are presented in Section 4, followed by empirical comparisons in Section 5.

2. Related Work

Blind deblurring with a single image has been an active field with many new methods emerging recently with different sparse image priors [6, 14, 7, 9, 20, 2]. In contrast, Rav-Acha and Peleg use two motion blurred images with different blur directions and show that restoration quality is superior than when using only a single image [12]. Since this work, many other multi-image blind deblurring algorithms have been developed [4, 3, 13, 22]. Most of these assume that only two blurry observations $\{\mathbf{y}_1, \mathbf{y}_2\}$ are present. In addition to other standard regularizers common to single-image blind deconvolution algorithms, a 'cross-blur' penalty function given by

$$E(\mathbf{k}_1, \mathbf{k}_2) = \|\mathbf{y}_1 * \mathbf{k}_2 - \mathbf{y}_2 * \mathbf{k}_1\|_2^2, \tag{2}$$

is often included [4, 13]. The rationale here is that, given the convolutional model from (1), $E(\mathbf{k}_1, \mathbf{k}_2)$ should be nearly zero if the noise levels are low and the correct kernels have been estimated. This penalty also implicitly relies on the coprimeness assumption, meaning that the blur kernels can only share a scalar constant [13]. Once the unknown kernels are estimated, the sharp image $\mathbf{x}$ may be recovered using a separate non-blind step if necessary.

Although computationally efficient, inclusion of this quadratic energy term does not always produce good kernel estimation results [4, 22]. One reason is that if the noise level is relatively high, it can dominate the minimization of $E(\mathbf{k}_1, \mathbf{k}_2)$, leading to kernel estimates that are themselves blurry, which may then either produce ringing artifacts or lost detail in the deblurred image [22]. Another issue is solution ambiguity, meaning that for a given optimal solution $\{\tilde{\mathbf{k}}_1, \tilde{\mathbf{k}}_2\}$, there exists a family of solutions $\{\tilde{\mathbf{k}}_1 * \mathbf{h}, \tilde{\mathbf{k}}_2 * \mathbf{h}\}$ that also minimize (2) [4, 22]. Finally, a practical limitation of $E(\mathbf{k}_1, \mathbf{k}_2)$ is that it only applies to images pairs, and hence would expand combinatorially as the number of observations grows.

To mitigate some of these problems, a sparse penalty on the blur kernel may be integrated into the estimation objective directly [4, 13] or applied via post-processing [22]. For example, Chen *et al.* propose a modified version of (2) that regularizes the kernel estimates using a sparse prior E_s, a continuity (smoothness) prior E_c, and a robust Lorentzian factor φ leading to the cost function

$$E(\mathbf{k}_1, \mathbf{k}_2) = \varphi(\mathbf{y}1 * \mathbf{k}_2 - \mathbf{y}_2 * \mathbf{k}_1) + \alpha \sum_{l=1}^{2} E_s(\mathbf{k}_l) + \beta \sum_{l=1}^{2} E_c(\mathbf{k}_l), \tag{3}$$

where α and β are trade-off parameters [4]. Similarly, Šroubek *et al.* modified (2) and incorporated a sparsity-promoting kernel prior based on a rectified ℓ_1-norm [13].

In contrast, Zhu *et al.* proposed a two-step approach for dual-image deblurring [22]. The blur kernels are first estimated using [13] with (2) incorporated. The resulting 'blurry' estimates $\{\hat{\mathbf{k}}_1, \hat{\mathbf{k}}_2\}$ are then refined in a second, sparsifying step. For this purpose, $\{\hat{\mathbf{k}}_1, \hat{\mathbf{k}}_2\}$ are treated as two blurry images whose sharp analogues are produced by minimizing

$$E(\mathbf{k}_1, \mathbf{k}_2, \mathbf{h}) = \sum_{l=1}^{2} \|\mathbf{k}_l * \mathbf{h} - \hat{\mathbf{k}}_i\|_2^2 + \alpha \sum_{l=1}^{2} \|\mathbf{k}_l\|_p^p \tag{4}$$

over $\mathbf{k}_1$, $\mathbf{k}_2$, and $\mathbf{h}$, with $p \leq 1$ producing a sparse ℓ_p norm over the kernels. This helps to remove the spurious factor $\mathbf{h}$ mentioned above while producing sparser kernel estimates. Although these approaches are all effective to some extent, the sparsity level of the blur kernels may require tuning.

In addition to motion-blurred observation processes, deblurring has also been attempted using images captured with different exposure lengths [18, 1] and coded apertures [21], etc. While using multiple images generally has the potential to outperform the single-image methods by fusing complementary information [12, 18, 4, 13], a principled approach that applies across a wide range of scenarios with little user-involvement or parameter tuning is still somewhat lacking. Our algorithm, which applies to any number of both noisy or blurry images without explicit trade-off parameters, is one attempt to fill this void.

3. Multi-Image Blind Deburring Algorithm

We will work in the derivative domain of images for ease of modeling and better performance [6, 9], meaning that $\mathbf{x}$ and $\mathbf{y}$ will denote the lexicographically ordered image derivatives of sharp and blurry images respectively obtained via a particular derivative filter.[1] Because convolution is a commutative operator, the blur kernels are unaltered.

Now consider the case where we have a single observation $\mathbf{y}$. The observation model from (1) defines a Gaussian likelihood function $p(\mathbf{y}|\mathbf{x}, \mathbf{k})$; however, maximum likelihood estimation of $\mathbf{x}$ and $\mathbf{k}$ is obviously ill-posed and hence

[1] The derivative filters used in this work are $\{[-1, 1], [-1, 1]^T\}$. Other choices are open.

we need a prior to regularize the solution space. In this regard, it is well-known that the gradients of sharp natural images tend to exhibit sparsity [10, 6, 9], meaning that many elements equal (or nearly equal) zero, while a few values remain large. With roots in convex analysis [11], it can be shown that essentially all iid distributions that favor such sparse solutions can be expressed as a maximization over zero-mean Gaussians with different variances. Mathematically, this implies that $p(\mathbf{x}) = \prod_{i=1}^{m} p(x_i)$ where m is the size of $\mathbf{x}$ ($\mathbf{y}$ is of size $n < m$) and

$$p(x_i) = \max_{\gamma_i \geq 0} \mathcal{N}(x_i; 0, \gamma_i) \exp\left[-\frac{1}{2} f(\gamma_i)\right]. \quad (5)$$

Here f is an arbitrary energy function. The hyperparameter variances $\boldsymbol{\gamma} = [\gamma_1, \ldots, \gamma_m]^T$ provide a convenient way of implementing several different estimation strategies [11]. For example, perhaps the most direct is a form of MAP estimation given by

$$\max_{\mathbf{x};\boldsymbol{\gamma},\mathbf{k} \geq 0} p(\mathbf{y}|\mathbf{x}, \mathbf{k}) \prod_i \mathcal{N}(x_i; 0, \gamma_i) \exp\left[-\frac{1}{2} f(\gamma_i)\right], \quad (6)$$

where simple update rules are available via coordinate ascent over $\mathbf{x}$, $\boldsymbol{\gamma}$, and $\mathbf{k}$ (a prior can also be included on $\mathbf{k}$ if desired). However, recently it has been argued that an alternative estimation procedure may be preferred for canonical sparse linear inverse problems [16]. The basic idea, which naturally extends to the blind deconvolution problem, is to first integrate out $\mathbf{x}$, and then optimize over $\mathbf{k}$, $\boldsymbol{\gamma}$, as well as the noise level λ. The final latent sharp image can then be recovered using the estimated kernel and noise level with standard non-blind deblurring algorithms.

Using the framework from [16], it can be shown that this alternative estimator is formally equivalent to solving

$$\min_{\mathbf{x};\mathbf{k},\lambda \geq 0} \frac{1}{\lambda} \|\mathbf{y} - \mathbf{k} * \mathbf{x}\|_2^2 + \mathfrak{g}(\mathbf{x}, \mathbf{k}, \lambda), \quad (7)$$

where $\mathfrak{g}(\mathbf{x}, \mathbf{k}, \lambda) \triangleq \min_{\boldsymbol{\gamma} \geq 0} \mathbf{x}^T \boldsymbol{\Gamma}^{-1} \mathbf{x} + \log|\lambda \mathbf{I} + \mathbf{H}\boldsymbol{\Gamma}\mathbf{H}^T|$, and $\mathbf{H}$ is the convolution matrix of $\mathbf{k}$. Note that this expression assumes that f is a constant; rigorous justification for this selection can be found in [17].

Optimization of (7) is difficult in part because of the high-dimensional determinants involved with realistic sized images. To alleviate this problem, we use determinant identities and a diagonal approximation to $\mathbf{H}^T\mathbf{H}$ as motivated in [9]. This leads to the simplified penalty function

$$\mathfrak{g}(\mathbf{x}, \mathbf{k}, \lambda) = \min_{\boldsymbol{\gamma} \geq 0} \sum_i \left[\frac{x_i^2}{\gamma_i} + \log(\lambda + \gamma_i \|\bar{\mathbf{k}}\|_2^2)\right], \quad (8)$$

where $\|\bar{\mathbf{k}}\|_2^2 \triangleq \sum_j k_j^2 \bar{I}_{ji}$ and $\bar{\mathbf{I}}$ is an indicator matrix with the j-th row recording the (column) positions where the j-th element of $\mathbf{k}$ appears in $\mathbf{H}$. $\|\bar{\mathbf{k}}\|_2^2$ can be viewed as the squared norm of $\mathbf{k}$ accounting for edge effects, or equivalently, as the squared norm of each respective column of $\mathbf{H}$. While technically then $\|\bar{\mathbf{k}}\|_2^2$ should depend on i, the column index of $\mathbf{H}$, we omit explicit referencing for simplicity.

In addition to many desirable attributes as described in [17], the cost function (7) provides a transparent entry-point for multi-image deblurring. Assuming that all observations $\mathbf{y}_l$ are blurry and/or noisy measurements of the same underlying image $\mathbf{x}$, then we may justifiably postulate that $\boldsymbol{\gamma}$ is shared across all l. This then leads to the revised, multi-image optimization problem

$$\min_{\mathbf{x},\{\mathbf{k}_l,\lambda_l \geq 0\}} \sum_{l=1}^{L} \frac{1}{\lambda_l} \|\mathbf{y}_l - \mathbf{k}_l * \mathbf{x}\|_2^2 + \mathfrak{g}(\mathbf{x}, \{\mathbf{k}_l, \lambda_l\}), \quad (9)$$

where the multi-image penalty function is now naturally defined as

$$\mathfrak{g}(\mathbf{x}, \{\mathbf{k}_l, \lambda_l\}) \triangleq \min_{\boldsymbol{\gamma} \geq 0} \sum_{l=1}^{L} \sum_{i=1}^{m} \left[\frac{x_i^2}{\gamma_i} + \log(\lambda_l + \gamma_i \|\bar{\mathbf{k}}_l\|_2^2)\right]. \quad (10)$$

The proposed cost function (9) can be minimized using coordinate descent (similar to MAP) outfitted with convenient upper bounds that decouple the terms embedded in (10). The resulting majorization-minimization approach, which is summarized in Algorithm 1, is guaranteed to reduce or leave unchanged (9) at each iteration, with similar convergence properties to the EM algorithm. Detailed derivations and connections with existing methods (e.g., Levin *et al.* [9]) can be found in [19].

While admittedly simple, the proposed model has a number of desirable features:

- It can handle a flexible number of degraded observations without requiring an extra 'cross-blurring' term, which generally limits the number of observations.
- The input can be a set of blurry or noisy observations without specifying the degradation type of each example; the algorithm will automatically estimate the blur kernel and the noise level for each one. We note that in the case of a single observation, the proposed method reduces to a robust single image blind deblurring model.
- The penalty function $\mathfrak{g}$ couples the latent image, blur kernels, and noise levels in a principled way. This leads to a number of interesting properties, including an inherent mechanism for scoring the relative quality each observed image during the recovery process and using this score to adaptively adjust the sparsity of the image regularizer. Section 4 is devoted to these developments.

Algorithm 1: Multi-Image Blind Deblurring.

Input: blurry images $\{\mathbf{y}_l\}_{l=1}^L$
Initialize: blur kernels $\{\mathbf{k}_l\}$, noise levels $\{\lambda_l\}$
While stopping criteria is not satisfied, do

- **Update** $\mathbf{x}$: $\mathbf{x} \leftarrow \left[\sum_{l=1}^L \frac{\mathbf{H}_l^T\mathbf{H}_l}{L\lambda_l} + \mathbf{\Gamma}^{-1}\right]^{-1} \sum_{l=1}^L \frac{\mathbf{H}_l^T\mathbf{y}_l}{L\lambda_l}$ where $\mathbf{H}_l$ is the convolution matrix of $\mathbf{k}_l$.
- **Update** $\boldsymbol{\gamma}$: $\gamma_i \leftarrow x_i{}^2 + \frac{\sum_{l=1}^L z_{li}}{L}$, $\mathbf{\Gamma} = \text{diag}(\boldsymbol{\gamma})$, $z_{li} \triangleq ((\sum_j k_{lj}^2 \bar{I}_{ji})\lambda_l^{-1} + \gamma_i^{-1})^{-1}$
- **Update** $\mathbf{k}_l$: $\mathbf{k}_l \leftarrow \arg\min_{\mathbf{k}_l \geq 0} \frac{1}{\lambda_l}\|\mathbf{y}_l - \mathbf{W}\mathbf{k}_l\|_2^2 + \sum_j k_{lj}^2(\sum_i z_{li}\bar{I}_{ji})$, with $\mathbf{W}$ the convolution matrix of $\mathbf{x}$
- **Update noise levels** λ_l: $\lambda_l \leftarrow \frac{\|\mathbf{y}_l - \mathbf{x}*\mathbf{k}_l\|_2^2 + \sum_{i=1}^m \sum_j k_{lj}^2 z_{li}\bar{I}_{ji}}{n}$

End

- The resulting algorithm (see Algorithm 1) is parameter-free thus requires minimal user involvement.

4. Penalty Function Analysis

This section will examine theoretical properties of the penalty function (10). These properties help to explain the success of our algorithm and hopefully demystify, at least to some extent, what otherwise may appear to be a somewhat non-standard, coupled regularizer that differs substantially from typical MAP estimators. For convenience, we first define

$$h(x, \boldsymbol{\rho}) \triangleq \min_{\gamma \geq 0} \sum_{l=1}^L \left[\frac{x^2}{\gamma} + \log(\rho_l + \gamma)\right]. \quad (11)$$

where $\boldsymbol{\rho} \triangleq [\rho_1, \ldots, \rho_L]^T$ with $\rho_l \triangleq \lambda_l / \|\bar{\mathbf{k}}_l\|_2^2$.[2] Then by noting the separability across pixels, (10) can be re-expressed as

$$\mathfrak{g}(\mathbf{x}, \{\mathbf{k}_l, \lambda_l\}) = \sum_{i=1}^m h(x_i, \boldsymbol{\rho}) + m\sum_{l=1}^L \log \|\bar{\mathbf{k}}_l\|_2^2, \quad (12)$$

which partitions image and kernel penalties into a more familiar form. The second term in (12) is similar to many common kernel penalties in the literature, and we will not consider it further here. However, the image penalty $h(x, \boldsymbol{\rho})$ is quite unique and we evaluate some of its relevant properties via two Theorems below followed by further discussion and analysis.

[2]Because of boundary effects, technically $\boldsymbol{\rho}$ will depend on i; however we omit this dependency to simplify notation.

Theorem 1 (Concavity) *The penalty function $h(x, \boldsymbol{\rho})$ is a concave non-decreasing function of $|x|$.*

Proofs will be deferred to [19]. Theorem 1 explicitly stipulates that a strong, sparsity promoting $\mathbf{x}$ penalty is produced by our framework, since concavity with respect to coefficient magnitudes is a well-known, signature property of sparse penalties [16]. Yet while this attribute may anchor our approach as a legitimate sparse estimator in the image (filter) domain, it does not explain precisely why it often produces superior results compared to more traditional MAP (or penalized regression) approaches, which also frequently possess a similar attribute (e.g., ℓ_1 norm-based penalties). For this purpose we must look deeper and examine how $\boldsymbol{\rho}$ modulates the effective penalty on $\mathbf{x}$.

First, for two values of the vector $\boldsymbol{\rho}$, e.g., $\boldsymbol{\rho}^1$ and $\boldsymbol{\rho}^2$, we use $\boldsymbol{\rho}^2 \succ \boldsymbol{\rho}^1$ to denote elementwise '$\geq$' with at least one element where the inequality is strict. We also define the function $h_{\boldsymbol{\rho}^\alpha} : \mathbb{R}^+ \rightarrow \mathbb{R}$ as $h_{\boldsymbol{\rho}^\alpha}(z) = h(z, \boldsymbol{\rho} = \boldsymbol{\rho}^\alpha)$, with domain $z \geq 0$. Note that because h is a symmetric function with respect to the origin, we may conveniently examine its concavity/curvature properties considering only the positive half of the real line.

Theorem 2 (Relative Sparsity) *The penalty function $h(x, \boldsymbol{\rho})$ is such that:*

1. *For all $\boldsymbol{\rho}^1$ and $\boldsymbol{\rho}^2$, $h_{\boldsymbol{\rho}^2}(z) - h_{\boldsymbol{\rho}^1}(z) \rightarrow 0$ as $z \rightarrow \infty$. Therefore, $h_{\boldsymbol{\rho}^1}$ and $h_{\boldsymbol{\rho}^2}$ penalize large magnitudes of x equally.*
2. *Let $\boldsymbol{\rho}^2 \succ \boldsymbol{\rho}^1$. Then if $z < z'$, we have $h_{\boldsymbol{\rho}^2}(z) - h_{\boldsymbol{\rho}^1}(z) > h_{\boldsymbol{\rho}^2}(z') - h_{\boldsymbol{\rho}^1}(z')$. Therefore, as $z \rightarrow 0$, $h_{\boldsymbol{\rho}^2}(z) - h_{\boldsymbol{\rho}^1}(z)$ is maximized, implying that $h_{\boldsymbol{\rho}^1}$ favors zero-valued coefficients more heavily than $h_{\boldsymbol{\rho}^2}$.*

From a more intuitive standpoint, $\boldsymbol{\rho}$ represents a form of shape parameter that modulates the concavity, or sparsity favorability, of the image prior. Moreover, each element of $\boldsymbol{\rho}$ can be viewed as a measure of the relative image estimation difficulty, with larger values indicative of greater difficulty. This is justified by the fact that larger values of some λ_l (meaning a higher noise level), or small values of some $\|\bar{\mathbf{k}}_l\|_2^2$ (meaning a more difficult, distributed kernel[3]), imply that $\rho_l = \lambda_l / \|\bar{\mathbf{k}}_l\|_2^2$ will be large.

More difficult cases (elements of $\boldsymbol{\rho}$ are large) occur for one of two reasons: (i) Either the underlying images are really corrupted by complex, diffuse blur kernels and/or high noise, or (ii) in the initial stages the algorithm has not been able to converge to a desirable, low-noise, low-blur solution. In both cases, the penalty function shape becomes nearly convex, which is highly desirable because it avoids

[3]For a given value of $\sum_i k_{li}$, a delta kernel maximizes $\|\bar{\mathbf{k}}_l\|_2^2$, while a kernel with equal-valued elements provides the minimum.

premature convergence to potentially suboptimal local solutions allowing coarse structures to be identified accurately.

In contrast, for cases where at least one image has a small ρ_l value, the effective penalty on $\mathbf{x}$ magnitudes becomes highly concave (sparsity favoring), even approaching a scaled (approximate) version of the ℓ_0 norm (in the sense described in [17]). This is because the existence of a single good kernel/noise estimation pair (meaning the associated ρ_l is small) necessitates that in all likelihood a good overall solution is nearby (even if some blur kernel/noise pairs associated with other observations are large). Fortunately, the $\log(\gamma + \rho_l)$ term associated with the l-th image will dominate the variational formation of $h(x, \boldsymbol{\rho})$, and a highly sparse, concave penalty function on $\mathbf{x}$ will ensue, allowing fine-grained kernel structures to be resolved. But there is now relatively little danger of local minima since we necessarily must be in the neighborhood of a good solution. The shape-adaptiveness of the coupled penalty function is the key factor leading the algorithm to success. Both the noise and blur dependency allow the algorithm to naturally possess a 'coarse-to-fine' estimation strategy, recovering large scale structures using a less aggressive (more convex) sparse penalty in the beginning, while later increasing its aggressiveness for recovering the small details. In so doing it can largely avoid being trapped in a local minima while recovering the blur kernel progressively. Please see [17, 19] for more details regarding why additional image sparsity is sometimes needed, as well as more information about how the proposed penalty function operates, including why it sometimes favors sparse kernels.

Interestingly, one auxiliary benefit of this procedure is that, given a set of corrupted image observations, and provided that at least one of them is reasonably good, the existence of other more highly degraded observations should not in theory present a significant disruption to the algorithm. In principle, such images are effectively discounted by the algorithm, largely as a consequence of Theorem 2, and each estimated ρ_l can be treated as a score function (see [19] for an empirical example and more details).

Finally, there is also a desirable form of scale invariance attributable to the proposed cost function, meaning that if $\mathbf{x}^*$ and $\{\mathbf{k}_l^*\}$ represent the optimal solution to (9) under the constraint $\sum_j k_{lj} = 1, \forall l$, then $\alpha^{-1}\mathbf{x}^*$ and $\{\alpha\mathbf{k}_l^*\}$ will always represent the optimal solution under the modified constraint $\sum_j k_{lj} = \alpha, \forall l$. Many previous models lack this type of scale invariance, and the exact calibration of the constraint (or related trade-off parameters) can fundamentally alter the form of the optimal solution beyond an irrelevant rescaling, thus require additional tuning.

5. Experimental Results

Using both synthetic data and real-world images, we now compare our algorithm with several *state-of-the-art* multi-image methods from Cai *et al.* [3], Šroubek *et al.* [13], and Zhu *et al.* [22] for blurry observations as well as Yuan *et al.* [18] and Whyte *et al.* [15] on noisy/blurry pairs.

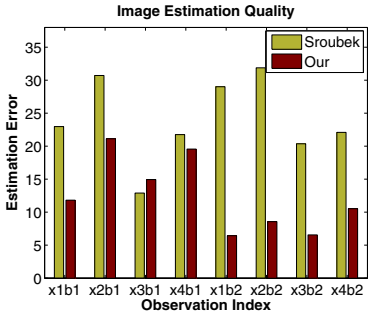

Figure 2. Error bar plot: Comparison of Šroubek *et al.*'s method [13] and ours on Levin *et al.*'s dataset [8].

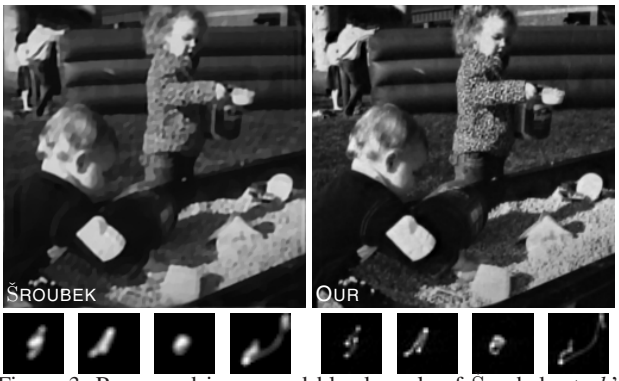

Figure 3. Recovered image and blur kernels of Šroubek *et al.*'s method [13] and ours on $\{x1, b1\}$, *i.e.*, the first image and kernels 1–4 from Levin *et al.*'s dataset [8].

5.1. Evaluation on Synthetic Data

We first use the standard test data collected by Levin *et al.* [8] for evaluation, which consists of 4 images of size 255×255 and 8 different blur kernels, giving a total of 32 blurry images. The kernel sizes range from 13×13 to 27×27. The blurry images, ground-truth images, and the ground-truth kernels are also provided. Following the experimental settings in [13], we construct a multi-observation test set with $L = 4$ blurry images by dividing the whole kernel set into two halves: $b1 = \{1 \cdots 4\}$ and $b2 = \{5 \cdots 8\}\}$. In so doing, 8 multi-observation sets are generated for testing. We then perform blind deblurring using different algorithms on each set. We compare our method with the recent method of Šroubek *et al.* [13], for which the `matlab` implementation is publicly available.[4]

The Sum of Squared Difference (SSD) metric defined in [8] is used for measuring the error between the deblurred and ground-truth images. Results are shown in Figure 2, where the proposed method generates deblurring results

[4] http://zoi.utia.cas.cz/files/fastMBD.zip

Figure 4. Dual motion deblurring results: (a) Blurry image pairs [3]. (b) Results from Cai *et al.* [3]. (c) Results produced with Šroubek *et al.*'s software [13]. (d) Our results.

that are significantly better on most of the test images. The recovered image and blur kernels from both methods for the first test set are shown in Figure 3. Here we observe that the kernels recovered by Šroubek *et al.* are overly blurry, thus leading to inferior image restoration quality (e.g., sand and sweater textures are compromised). In contrast, our approach can recover the blur kernels with high quality *without* using any explicit sparse prior over the kernel. By incorporating a sparsity prior over the kernel, the results can be further improved (results not shown). Overall, the more refined kernel estimates obtained via the proposed approach translate into more details recovered in the latent images.

5.2. Evaluation on Real-World Images

Blind restoration using multiple observations is a ubiquitous problem, with many potential applications. This section investigates two common scenarios using real-world images: *dual motion deblurring*, *i.e.*, using two motion-blurried observations for joint blind deblurring [12, 3, 4, 13, 22]; and *blurry/noisy pair restoration*, *i.e.*, using a short-exposure noisy and long-exposure blurry image pair for joint restoration [18, 15]. We emphasis that the reason we evaluate under these somewhat restrictive scenarios separately is primarily for ease of comparison with previous state-of-the-art algorithms that have been explicitly tailored for each specific case. In contrast, our algorithm does not require any modification and can handle both tasks seamlessly in a unified way, and is in this sense more practical.

Dual Motion Deblurring For dual motion deblurring, we compare with the multi-image methods proposed by Cai *et al.* [3], Šroubek *et al.* [13] as well as Zhu *et al.* [22] on several different real-world images used in previous deblurring work. We first evaluate the relative performance on an image pair from [3] as shown in Figure 4. The results of Cai *et al.*, Šroubek *et al.*, and our method are also shown in Figure 4, with the estimated blur kernels displayed in the top-right corner of each image. We observe that the kernel estimates produced by Šroubek *et al.* may be overly diffuse, at least not precise enough for generating a crisp deblurring result with limited ringing. The kernel estimations from Cai *et al.* are better, but the recovered image still suffers from some ringing artifacts. While we do not have access to the ground-truth kernel for real-world images, the relatively compact support of our kernel appears to be reasonable given the high quality of the estimated sharp image.

Figure 5 provides further comparison with Šroubek *et al.* on an image pair from [13], as well as with a standard single-image method from Cho *et al.* [5] as a benchmark. The kernel estimates from Šroubek *et al.* may again be overly diffuse, and the associated latent image estimate retains some significant artifacts.

The method of Zhu *et al.* [22] attempts to refine the estimated blur kernels from Šroubek *et al.*[13] via an explicit sparsity penalty. Comparisons between Zhu *et al.*, Šroubek *et al.*, and our approach on an image pair from [22] are shown in Figure 6. While the kernel estimates from Zhu *et al.* are indeed more compact than those from Šroubek *et al.*, the accuracy is likely still below that of our method. For example, some fine details such as the text on the book cover are not properly recovered. One potential reason for this is that the kernel refining step of Zhu *et al.* relies purely on the kernels estimated via Šroubek *et al.*, without using the observed data. Therefore, although the estimated blur kernels do become less diffuse, they are not necessarily consistent with the observed data, as any error generated in the original kernel estimation step will be inevitably transferred during the kernel refining process. In contrast, our approach can implicitly determine the proper kernel sparsity directly from the data without any secondary rectifications or an explicit sparse prior for the kernel; it therefore appears to be more reliable on these test images.

Restoration from Blurry/Noisy Pairs As mentioned previously, our algorithm can be seamlessly applied to images with differing types of degradation extending beyond the

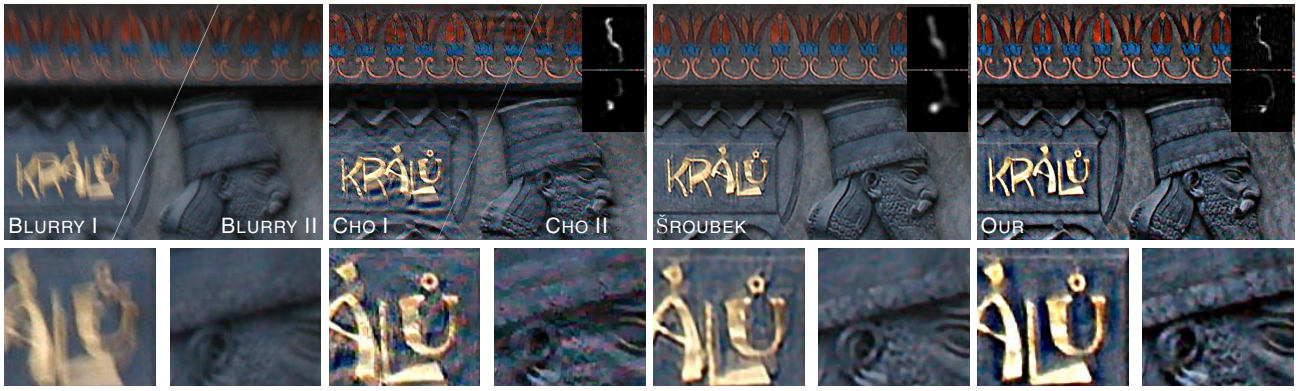

Figure 5. Dual motion deblurring results: (a) Blurry image pair [13]. (b) Results produced with Cho *et al.*'s software [5]. (c) Results produced with Šroubek *et al.*'s software [13]. (d) Our results.

Figure 6. Dual motion deblurring results: (a) Blurry image pair [22]. (b) Results produced with Šroubek *et al.*'s software [13]. (c) Results from Zhu *et al.* [22]. (d) Our results.

typical dual-motion deblurring tasks, e.g., restoration based on blurry/noisy pairs. Although the existing dual-motion deblurring algorithms tested above are no longer directly applicable, alternative approaches have been specifically tailored to work only with a blurry and noisy pair [18, 15], and hence provide a benchmark for comparison.

We first compare with Yuan *et al.* [18] on the blurry/noisy image pair previously used in their paper. The results, including Cho *et al.*'s method [5] as a single-blurry-image baseline, are shown in Figure 7. Not surprisingly, Yuan *et al.* can generate a restoration result that is of higher quality compared to the result obtained from a single blurry image and Cho *et al.*'s algorithm. Yet the image recovered via our approach is of relatively similar quality to that of Yuan *et al.*; however, we emphasize that our method is at a substantial disadvantage because it has no knowledge that we are dealing with a blurry/noisy pair and it has received no special design for this situation. It is also interesting to point out that the blur kernel estimated for the noisy image is a delta kernel as would be expected if the correct solution were to be found. This reflects the strong generalization ability of our method.

Finally, we compare with two recent blurry/noisy pair-based methods from Whyte *et al.* [15] using images from their paper. Results are shown in Figure 8. Note that Whyte *et al.*'s non-uniform method does not produce a typical 2D kernel per the standard convolutional model (1), and hence no blur kernel is shown. Again, we observe that our algorithm, without resorting to more complicated observation models or special tuning, performs competitively with algorithms specifically designed to work with a known blurry and noisy pair.

6. Conclusion

By utilizing a novel penalty function that couples the latent sharp image, blur kernels, and noise variances in a theoretically well-motivated way, this paper describes a unified multi-image blind deconvolution algorithm applicable for recovering a latent, high-quality image from a given set of degraded (blurry, noisy) observations, without any specific modifications for different types of degradations. Moreover, it automatically adapts to the quality of each observed image, allowing higher quality images to dominate the estimation process when appropriate. Experimental evaluations validate the proposed method in different multi-image restoration scenarios. For future work, we would like to generalize our algorithm to areas such as video deblurring and non-uniform deblurring. At least for single images, we have already found that our method performs well with non-uniform camera shake provided appropriate basis functions

Figure 7. Dual exposure deblurring results: (a) Blurry/Noisy image pair [18]. (b) Results produced with Cho *et al.*'s software using the blurry image [5]. (c) Result from Yuan *et al.* [18]. (d) Our results.

Figure 8. Dual exposure deblurring results: (a) Blurry/Noisy image pair [15]. (b) Uniform deblurring results from Whyte *et al.* [15]. (c) Non-uniform deblurring result from Whyte *et al.* [15]. (d) Our results.

are adopted for representing $\mathbf{H}$.

Acknowledgements This work was done while the first author was an intern at Microsoft Research Asia. It was also supported by the Excellent Doctorate Foundation of NPU and the Ministry of Education Fund for Doctoral Students Newcomer Awards and NSF of China (61231016).

References

[1] A. K. Agrawal, Y. Xu, and R. Raskar. Invertible motion blur in video. *ACM Trans. Graph.*, 28(3), 2009. 2

[2] S. D. Babacan, R. Molina, M. N. Do, and A. K. Katsaggelos. Bayesian blind deconvolution with general sparse image priors. In *ECCV*, 2012. 2

[3] J.-F. Cai, H. Ji, C. Liu, and Z. Shen. Blind motion deblurring using multiple images. *J. Comput. Physics*, 228(14):5057–5071, 2009. 1, 2, 5, 6

[4] J. Chen, L. Yuan, C.-K. Tang, and L. Quan. Robust dual motion deblurring. In *CVPR*, 2008. 1, 2, 6

[5] S. Cho and S. Lee. Fast motion deblurring. In *SIGGRAPH ASIA*, 2009. 6, 7, 8

[6] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman. Removing camera shake from a single photograph. In *SIGGRAPH*, 2006. 1, 2, 3

[7] D. Krishnan, T. Tay, and R. Fergus. Blind deconvolution using a normalized sparsity measure. In *CVPR*, 2011. 2

[8] A. Levin, Y. Weiss, F. Durand, and W. Freeman. Understanding and evaluating blind deconvolution algorithms. In *CVPR*, 2009. 5

[9] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman. Efficient marginal likelihood optimization in blind deconvolution. In *CVPR*, 2011. 2, 3

[10] J. W. Miskin and D. J. C. MacKay. Ensemble learning for blind image separation and deconvolution. In *Advances in Independent Component Analysis*, 2000. 3

[11] J. A. Palmer, D. P. Wipf, K. Kreutz-Delgado, and B. D. Rao. Variational EM algorithms for non-gaussian latent variable models. In *NIPS*, 2006. 3

[12] A. Rav-Acha and S. Peleg. Two motion blurred images are better than one. *Pattern Recognition Letters*, 26:311–317, 2005. 1, 2, 6

[13] F. Šroubek and P. Milanfar. Robust multichannel blind deconvolution via fast alternating minimization. *IEEE Trans. on Image Processing*, 21(4):1687–1700, 2012. 1, 2, 5, 6, 7

[14] Q. Shan, J. Jia, and A. Agarwala. High-quality motion deblurring from a single image. In *SIGGRAPH*, 2008. 1, 2

[15] 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. 5, 6, 7, 8

[16] D. P. Wipf, B. D. Rao, and S. S. Nagarajan. Latent variable Bayesian models for promoting sparsity. *IEEE Trans. on Information Theory*, 57(9):6236–6255, 2011. 3, 4

[17] D. P. Wipf and H. Zhang. Revisiting Bayesian blind deconvolution. Technical report, MSRA, Feb. 2013. 3, 5

[18] L. Yuan, J. Sun, and H.-Y. Shum. Image deblurring with blurred/noisy image pairs. In *SIGGRAPH*, 2007. 2, 5, 6, 7, 8

[19] H. Zhang, D. P. Wipf, and Y. Zhang. Multi-observation blind deconvolution with an adaptive sparse prior. Technical report, MSRA, March 2013. 3, 4, 5

[20] H. Zhang, J. Yang, Y. Zhang, N. M. Nasrabadi, and T. S. Huang. Close the loop: Joint blind image restoration and recognition with sparse representation prior. In *ICCV*, 2011. 2

[21] C. Zhou, S. Lin, and S. Nayar. Coded aperture pairs for depth from defocus and defocus deblurring. *International Journal of Computer Vision*, 93(1):53–72, 2011. 2

[22] X. Zhu, F. Šroubek, and P. Milanfar. Deconvolving PSFs for a better motion deblurring using multiple images. In *ECCV*, 2012. 1, 2, 5, 6, 7