Why Maximum-A-Posteriori Blind Image Deblurring Works After All

Gian Marti, Boxiao Ma, and Hans-Andrea Loeliger
Dept. of Information Technology & Electrical Engineering
ETH Zurich
{marti, ma, loeliger}@isi.ee.ethz.ch

***Abstract*—Maximum-a-posteriori (MAP) methods, while being a standard choice for many estimation problems, have been considered problematic for blind image deblurring: They have been suspected of preferring blurry images to sharp ones. Alternative methods without this apparent defect have been proposed instead. Reservations about MAP methods for blind image deblurring persist even as their close relation to these alternatives has become evident.**

We revisit the literature on this topic and argue that the original rejection of MAP methods was ill-founded. We show that the MAP approach can prefer sharp images over blurry ones. Furthermore, we show experimentally that the MAP approach can in principle achieve deblurring results that are competitive with the allegedly superior methods. We thereby challenge some traditional notions of the relevant causes underlying successful blind deblurring to obtain a more accurate understanding of the blind image deblurring problem.

***Index Terms*—Blind image deblurring, MAP, sparsity, smoothed NUV, iteratively reweighted coordinate descent.**

I. Introduction

Recovering a sharp image from a single blurry observation is an important problem in image processing with a long history [1]–[3]. The relation between the blurry observation and the assumed sharp underlying image is often [2], [3] modeled as

$$\mathbf{y} = \mathbf{x} * \mathbf{k} + \mathbf{n} \tag{1}$$

where $\mathbf{y}$ is the blurry image, $\mathbf{x}$ is the sharp underlying image, $\mathbf{k}$ is the blur-kernel, and $\mathbf{n}$ is observation noise. In this paper, it is assumed that the kernel $\mathbf{k}$ is invariant over the whole image. The task is to estimate $\mathbf{x}$ from $\mathbf{y}$ alone. (Typically, estimating $\mathbf{x}$ involves also estimating $\mathbf{k}$.) This problem is ill-posed as there are infinitely many triples $(\mathbf{x}, \mathbf{k}, \mathbf{n})$ satisfying (1), and even infinitely many pairs $(\mathbf{x}, \mathbf{k})$ satisfying $\mathbf{y} = \mathbf{x} * \mathbf{k}$. It is therefore not immediately evident how one should tackle the problem in order to arrive at a satisfactory solution. A straightforward approach is to set up a statistical model and to use a maximum-a-posteriori (MAP) estimator for finding a pair $(\hat{\mathbf{x}}, \hat{\mathbf{k}})$ that maximizes the posterior

$$p(\mathbf{x}, \mathbf{k}|\mathbf{y}) \propto p(\mathbf{y}|\mathbf{x}, \mathbf{k})\, p(\mathbf{x})\, p(\mathbf{k}). \tag{2}$$

The observation model $p(\mathbf{y}|\mathbf{x}, \mathbf{k})$ is chosen according to (1) where the noise $\mathbf{n}$ is Gaussian, $\mathbf{n} \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma_n^2)$. The image prior $p(\mathbf{x})$ is a factorizing sparse prior on the image gradients, as natural images are known to be approximately sparse in the gradient domain [4]. The kernel prior $p(\mathbf{k})$ is taken to be constant over the set $\{\mathbf{k} \geq 0,\ \|\mathbf{k}\|_1 = 1\}$. Taking the negative logarithm of (2) yields the equivalent optimization problem

$$\underset{(\mathbf{x},\mathbf{k}):\, \mathbf{k}\geq 0,\, \|\mathbf{k}\|_1 = 1}{\arg\min} L(\mathbf{x}, \mathbf{k}), \tag{3}$$

where the cost function to be minimized is

$$L(\mathbf{x}, \mathbf{k}) = \frac{\|\mathbf{y} - \mathbf{x} * \mathbf{k}\|^2}{2\sigma_n^2} + \sum_{(i,j)\in\Delta} \kappa(x_i - x_j). \tag{4}$$

Here, Δ denotes the set of neighbored pixels and the function $\kappa(\cdot)$ is induced by $p(\mathbf{x})$ through the relation $p(\mathbf{x}) = \prod_{(i,j)\in\Delta} \exp(-\kappa(x_i - x_j))$. We refer to this approach as *naive MAP*.

In a seminal paper [5], it was argued that naive MAP is bound to fail: It has been observed that sparse priors on the image gradients tend to favor blurry images over sharp ones, and it was claimed that this leads to systematic preference of the degenerate "no-blur" solution $(\hat{\mathbf{x}}, \hat{\mathbf{k}}) = (\mathbf{y}, \delta)$ over the "true" sharp image and blur kernel. As part of their analysis, the authors also made available a dataset which they used for illustrating this pathological behavior of the naive MAP [5]. However, the naive MAP's basic ability to achieve sensible deblurring results has been shown both in earlier experiments [6] and more recently [7], [8]. A variety of explanations were offered for the discrepancy between these empirical results and the analysis in [5]. Many of these accepted the conclusion that there are inherent problems with a naive MAP approach to blind image deblurring.

In this paper, we review these analyses: We argue that the original dismissal by [5] of naive MAP was unjustified, and that the naive MAP can, in fact, prefer sharp images over blurry ones. We also show empirically that naive MAP can—at least on the dataset [5] used for its original dismissal—perform on par with or better than many of the methods that have been designed as a solution to its alleged defects.

Our contributions are the following: We provide a brief but cohesive review of the literature on blind MAP deblurring. Second, we add our own analysis in an attempt to resolve the debate surrounding the naive MAP approach to blind image deblurring. Finally, we evaluate our own naive MAP method empirically on the original dataset [5]. We find that the results of the method are at least on par with or better than any of

the discussed approaches which attempt to avoid the MAP's perceived problems. Our overall conclusion is not that naive MAP is the best approach to blind image deblurring—concerns about robustness and computational efficiency, for instance, are completely bypassed in our analysis. Rather, our overall conclusion will be that (with respect to the model (1)) the naive MAP is able not only to avoid the no-blur solution but to deliver highly accurate deblurring results. We believe it is of conceptual interest that the naive MAP approach to blind image deblurring is not flawed. Indeed, a misunderstanding of this point may indicate an insufficient understanding of the blind deblurring problem.

The paper is structured as follows: In Section II, we revisit in a cohesive manner some of the literature on the MAP for blind image deblurring, adding our own analysis. In Section III, we provide an empirical comparison of our own naive MAP approach to the methods discussed in Section II. This empirical comparison is performed on the original dataset of [5]. Finally, in Section IV, we make some conclusions.

In the remainder of the paper, we assume that the blurry image $\mathbf{y}$ has been generated by a ground-truth kernel $\mathbf{k}_0$ and sharp image $\mathbf{x}_0$, i.e., $\mathbf{y} = \mathbf{x}_0 * \mathbf{k}_0$. (The dataset of [5] fulfills this assumption.) We denote by $\mathbf{x}_{\mathbf{k}_0}$ and $\mathbf{x}_\delta$ the minimizers of (3) with respect to $\mathbf{x}$ while holding $\mathbf{k}$ fixed to $\mathbf{k}_0$ and δ (the Kronecker delta), respectively.

II. Analyzing the Maximum-a-Posteriori Approach to Blind Image Deblurring

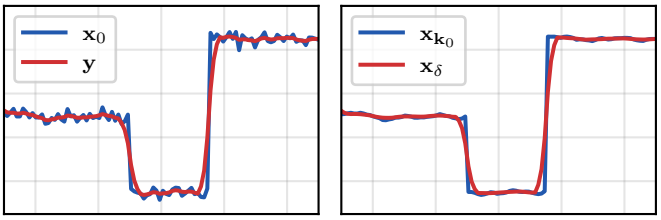
A. Reviewing the Literature

Different image estimates $\mathbf{x}$ can, by choice of a suitable kernel $\mathbf{k}$, yield identical noise-free blurry images $\mathbf{x} * \mathbf{k}$ and hence yield the same likelihood $p(\mathbf{y}|\mathbf{x}, \mathbf{k})$. As $p(\mathbf{k})$ is constant over its domain, the naive MAP relies exclusively on the image prior $p(\mathbf{x})$ to distinguish between such candidate estimates. When comparing the true explanation $(\mathbf{x}_0, \mathbf{k}_0)$ with the no-blur explanation $(\mathbf{y}, \delta)$, where the kernel estimate is a Kronecker delta and the image estimate equals the blurry image, it has been recognized in [5] that there are two opposing effects at work:

- The cost function $\kappa(\cdot)$ induced by a sparse prior is typically concave and therefore prefers sharp pixel differences over blurred-out ones, favoring the true explanation $\mathbf{x}_0$ over the blurry explanation $\mathbf{y}$.
- However, $\kappa(\cdot)$ is typically also strictly increasing, so flat surfaces are preferred over textured ones, favoring the no-blur explanation over the true one.

Fig. 1(a) depicts a schematic 1D example: $\mathbf{x}_0$ is preferred for its clear-cut edges while $\mathbf{y}$ is preferred for its flatter surfaces. The key finding of [5] was that, on natural images, the detrimental effect consistently dominates the benign one, implying that $L(\mathbf{y}, \delta) < L(\mathbf{x}_0, \mathbf{k}_0)$. This is the case even if $p(\mathbf{x})$ accurately reflects the statistics underlying the sharp image $\mathbf{x}_0$.

The widely accepted conclusion drawn in [5] was that the naive MAP is inherently unsuited for blind image deblurring, and alternative approaches that would avoid this alleged degeneracy were pursued instead.



(a) 1D toy example of a sharp signal $\mathbf{x}_0$ and a blurred signal $\mathbf{y}$. (b) The minimizers of (4) when $\mathbf{k}$ is fixed to $\mathbf{k}_0$ and δ, respectively.

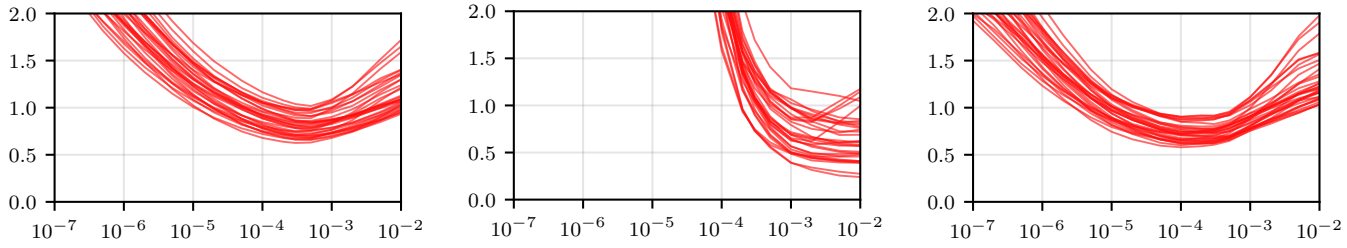
Fig. 1: Schematic comparison of blurred and sharp signals. Left: Blurry signal and sharp ground truth. Right: Smoothed signals when minimizing (4) in $\mathbf{x}$ for the given $\mathbf{y}$ and kernels $\mathbf{k}_0$ and δ, respectively.

Based on the different dimensionalities of $\mathbf{x}$ and $\mathbf{k}$, the following strategy was advocated in [5] itself: Integrating out $\mathbf{x}$ in (2) to estimate the kernel as $\hat{\mathbf{k}} = \arg\max p(\mathbf{k}|\mathbf{y})$, and subsequently estimating the image as $\hat{\mathbf{x}} = \arg\max p(\mathbf{x}|\hat{\mathbf{k}}, \mathbf{y})$. In order to approximate the intractable integration, it was suggested to follow Fergus et al. [9] in using a variational Bayesian (VB) strategy and a diagonal approximation to a certain covariance matrix. Such a VB strategy was also employed by Babacan et al. [10].

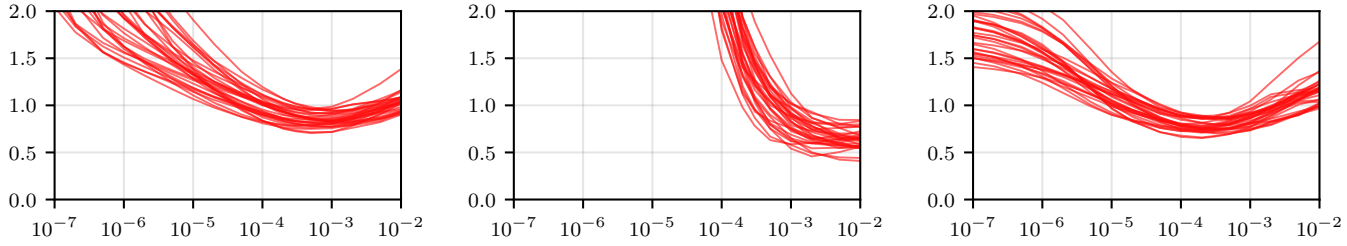
Alternatively, Xu and Jia [11], and Xu et al. [12] proposed to estimate the kernel from latent images that contain only the salient structures of $\mathbf{x}_0$. These salient structures were either detected explicitly [11] or brought out through "L_0-regularization" of the image gradients [12].

All the while, however, the naive MAP has not been a total failure; e.g., Chan and Wong [6] reported successful results using total variation regularization. Perrone and Favoro [7] noted this and attempted to reconcile it with the analysis of [5]. They argued that this suprising success is effected by an algorithmic peculiarity of [6] (delaying the projection of the kernel estimate onto the feasible set when optimizing). Effectively, it was claimed, the deblurring algorithm of [6] would optimize an objective which differs from the naive MAP objective (4). The argument of [7] rests on a theoretical analysis of certain 1D toy examples for which the authors were able to prove the pertinence of the distinction between the deblurring objectives. They also showed empirically that a modified MAP-approach as in [6] can produce respectable results on the dataset of [5]. However, it is not clear that the difference between the objectives is as decisive for natural image deblurring as it is for the 1D toy signals considered in [7].

In contrast, Wipf and Zhang [13] showed that the VB strategy advocated in [5] becomes equivalent to a MAP-approach with an unorthodox image prior once the necessary approximations for tractability are made as in [9], [10]. They argued that the special characteristics of the resultant image prior (which introduces an interdependence between $\mathbf{x}, \mathbf{k}$ and σ_n^2) are responsible for the empirical superiority of VB strategies over the naive MAP. Still, this fundamental relation between the successful VB strategies and the MAP called into question the latter's inherent unsuitedness for blind



(a) $L(\breve{\mathbf{x}}_{\mathbf{k}_0}, \mathbf{k}_0)/L(\breve{\mathbf{x}}_\delta, \delta)$ (b) $\|\mathbf{y} - \breve{\mathbf{x}}_{\mathbf{k}_0} * \mathbf{k}_0\|^2 / \|\mathbf{y} - \breve{\mathbf{x}}_\delta * \delta\|^2$ (c) $\kappa(\breve{\mathbf{x}}_{\mathbf{k}_0})/\kappa(\breve{\mathbf{x}}_\delta)$



(d) $L(\breve{\mathbf{x}}^{\downarrow}_{\mathbf{k}_0}, \mathbf{k}^{\downarrow}_0)/L(\breve{\mathbf{x}}^{\downarrow}_\delta, \delta)$ (e) $\|\mathbf{y}^{\downarrow} - \breve{\mathbf{x}}^{\downarrow}_{\mathbf{k}_0} * \mathbf{k}^{\downarrow}_0\|^2 / \|\mathbf{y}^{\downarrow} - \breve{\mathbf{x}}^{\downarrow}_\delta * \delta\|^2$ (f) $\kappa(\breve{\mathbf{x}}^{\downarrow}_{\mathbf{k}_0})/\kappa(\breve{\mathbf{x}}^{\downarrow}_\delta)$

Fig. 2: Top row: Cost ratios between the estimates $\breve{\mathbf{x}}_{\mathbf{k}}$ and $\breve{\mathbf{x}}_\delta$ as functions of σ_n^2 (a), as well as the ratios between the respectively incurred costs of the data fit term (b) and the prior term (c). Bottom row: Analog quantities for the downscaled problem instance.

image deblurring. Wipf and Zhang [13] concluded that the essence of a good prior $p(\mathbf{x})$ for blind image deblurring lies in the following property: For some visually convincing $\hat{\mathbf{x}}$ which in combination with $\mathbf{k}_0$ satisfactorily explains $\mathbf{y}$ (i.e., $\hat{\mathbf{x}} * \mathbf{k}_0 \approx \mathbf{x}_0 * \mathbf{k}_0$), the prior should satisfy $p(\mathbf{x}_0 * \mathbf{k}_0) \ll p(\hat{\mathbf{x}})$, meaning that it favors the desired explanation $(\hat{\mathbf{x}}, \mathbf{k}_0)$ over the "no-blur" explanation $(\mathbf{x}_0 * \mathbf{k}_0, \delta)$. They recognized the "$L_0$-prior" of [12] as such a prior but accepted the conclusion from [5] that natural image priors fall short in this regard and are unsuited for blind image deblurring.

This notion was finally challenged by Cho and Lee [8], who showed empirically that more natural image priors can also induce a cost function with the property $L(\mathbf{x}_{\mathbf{k}_0}, \mathbf{k}_0) < L(\mathbf{x}_\delta, \delta)$. The apparent contradiction between this finding and the analysis in [5] is resolved by noting that, even though $\mathbf{x}_0$ does have lower probability and therefore a higher incurred cost under the prior $p(\mathbf{x})$, $\mathbf{x}_{\mathbf{k}_0}$—being a minimizer of (4)—will be a smoothed version of $\mathbf{x}_0$ (see Fig. 1(b) for a schematic depiction) with a correspondingly higher probability under $p(\mathbf{x})$. Here, $\mathbf{x}_{\mathbf{k}_0}$ should be identified with $\hat{\mathbf{x}}$ from the preceding paragraph. Only in the limit $\sigma_n^2 \to 0$ is $\mathbf{x}_{\mathbf{k}_0}$ bound to equal $\mathbf{x}_0$ (and $\mathbf{x}_\delta$ bound to equal $\mathbf{y}$), so that the problematic relation $L(\mathbf{x}_{\mathbf{k}_0}, \mathbf{k}_0) > L(\mathbf{x}_\delta, \delta)$ holds.[1] Cho and Lee [8] also evaluated their own MAP approach on the dataset [5], finding it to consistently avoid the no-blur solution and yielding acceptable results.

[1]Actually, $\mathbf{x}_{\mathbf{k}_0}$ need not equal $\mathbf{x}_0$ even as $\sigma_n^2 \to 0$ since the mapping $\mathbf{x} \mapsto \mathbf{k}_0 * \mathbf{x}$ need in general not be injective. However, this qualifier is irrelevant to our argument.

B. Extending the analysis

Our own analysis resembles Cho and Lee's [8] and also focuses on the images of the dataset [5]. However, their analysis is entirely in the gradient domain, meaning that their cost function is of the form

$$\frac{\|\mathbf{y}' - \mathbf{x}' * \mathbf{k}\|^2}{2\sigma_n^2} + \sum_i \kappa(x'_i), \tag{5}$$

where $\mathbf{x}'$ and $\mathbf{y}'$ represent image gradients. This strategy is natural and common since convolution preserves the equality $\mathbf{y} = \mathbf{x} * \mathbf{k}$ when taking the gradients of $\mathbf{x}$ and $\mathbf{y}$. But the quantity $\|\mathbf{y} - \mathbf{x} * \mathbf{k}\|^2$ is not invariant to taking the gradients of $\mathbf{x}$ and $\mathbf{y}$. While this does not invalidate their results, the setup in [8] thereby deviates from the setup of [5].

In contrast, we consider a cost function of the form (4) to show that $L(\mathbf{x}_{\mathbf{k}_0}, \mathbf{k}_0)$ may be lower than $L(\mathbf{x}_\delta, \delta)$ also in the setup of [5].[2] We also go beyond [8] in analyzing the individual contributions of the data-fit term and the prior term to the total cost, and in analyzing how downscaling the images affects the relation between $L(\mathbf{x}_{\mathbf{k}_0}, \mathbf{k}_0)$ and $L(\mathbf{x}_\delta, \delta)$.[3] We use the cost function (4), with

$$\kappa(u) \triangleq \begin{cases} \log\left|\frac{u}{\sigma_0}\right| + \frac{1}{2} & \text{if } u^2 > \sigma_0^2 \\ \frac{u^2}{2\sigma_0^2} & \text{otherwise.} \end{cases} \tag{6}$$

[2]It might be objected that our function $\kappa(\cdot)$ in (6) is different from the $\kappa(\cdot) = |\cdot|^\alpha, 0 < \alpha < 1$ in [5]. While we agree that our $\kappa(\cdot)$ is better suited for the task, the analysis and findings of [5]—i.e., that $\kappa(\mathbf{x}_0) > \kappa(\mathbf{y})$—remain unchanged for our $\kappa(\cdot)$.

[3]Solving downscaled problem instances and using the solutions for initialization when solving the original problem is a widely used strategy in blind image deblurring.

This function $\kappa(\cdot)$, which is the *plain SNUV* function of [14], admits a variational representation in the sense of [15]. It has recently been used for multi-image blind deblurring [16]. For the experiments in this paper, the parameter σ_0^2 is set to 10^{-4}.

Ideally, we would assess the ratio $L(\mathbf{x}_{\mathbf{k}_0}, \mathbf{k}_0)/L(\mathbf{x}_\delta, \delta)$. As (4) is a non-convex problem, its global minimizer is difficult to obtain. We therefore resort to the proxy $L(\breve{\mathbf{x}}_{\mathbf{k}_0}, \mathbf{k}_0)/L(\breve{\mathbf{x}}_\delta, \delta)$, where $\breve{\mathbf{x}}_{\mathbf{k}_0}$ and $\breve{\mathbf{x}}_\delta$ are approximations to $\mathbf{x}_{\mathbf{k}_0}$ and $\mathbf{x}_\delta$ that are obtained by iteratively reweighted coordinate descent (IRCD) as in [16]. We maintain that this proxy is valid for two reasons: First, since the same algorithm is used for optimizing both problem instances, the estimate of the cost ratio should not be biased. Second, even if this estimate is biased, the relation $L(\breve{\mathbf{x}}_{\mathbf{k}_0}, \mathbf{k}_0)/L(\breve{\mathbf{x}}_\delta, \delta) < 1$ still indicates that the cost function favors the sharp explanation $(\breve{\mathbf{x}}_{\mathbf{k}_0}, \mathbf{k}_0)$ over the no-blur explanation $(\breve{\mathbf{x}}_\delta, \delta)$. So successful deblurring can be explained in this way, without arguing that the optimization algorithm systematically produces results which are at odds with the purported cost function (as is done in [7]).

The estimated ratios $L(\breve{\mathbf{x}}_{\mathbf{k}_0}, \mathbf{k}_0)/L(\breve{\mathbf{x}}_\delta, \delta)$ for the dataset [5] are depicted as functions of σ_n^2 in Fig. 2(a). Figs. 2(b) and 2(c) depict the ratios of the data-fit and prior term of (4), respectively. Note that for suitable choice of $\sigma_n^2 \in [10^{-4}, 10^{-3}]$, the cost function overwhelmingly favors the sharp explanation, and that this is primarily induced by the lower prior cost of $\breve{\mathbf{x}}_{\mathbf{k}_0}$. For very small σ_n^2, the effect described by [5] occurs with respect to the prior cost (Fig. 2(c)) and also the data-fit is much worse for $\breve{\mathbf{x}}_{\mathbf{k}_0}$ than for $\breve{\mathbf{x}}_\delta$ (Fig. 2(b)), so that the no-blur explanation is favored over the sharp one. In contrast, for very large σ_n^2, both $\breve{\mathbf{x}}_{\mathbf{k}_0}$ and $\breve{\mathbf{x}}_\delta$ are heavily smoothed images which need to account for $\mathbf{y}$ only loosely. There is no consistent discrimination between them in terms of cost in this regime.

The bottom row of Fig. 2 depicts analogous statistics for versions of $\mathbf{y}$ and $\mathbf{k}_0$ that have been downscaled by approximately a factor of 2 using bilinear interpolation. (We use the superscript "$\downarrow$" to denote the downsampled quantities, e.g. $\mathbf{y}^\downarrow$.) Note that the cost function now consistently favors the desired solution if $\sigma_n^2 \in [10^{-4}, 10^{-3}]$. This happens even though bilinear interpolation does not preserve the inequality $\mathbf{y} = \mathbf{k}_0 * \mathbf{x}_0$, so that $\mathbf{k}_0^\downarrow$ is at best a crude approximation to the true blur kernel at smaller scale. The implication is that naive MAP should also be applicable in combination with a multi-resolution scheme.

III. Empirical Evaluation

Even a vanilla variant of naive MAP deblurring using the cost function from Section II-B and IRCD [14], [16] for optimization is successful in blind deblurring. But optimal results are achieved when some common tricks of the trade are used:

First, a multi-resolution scheme is used, where $\mathbf{k}$ is initialized based on the kernel estimate from a downsampled version of the deblurring problem: Section II-B has shown that the discrimination between the desired and the no-blur solution is at least as good for the downsampled problem, and the search space is smaller, leading to more robust inference.

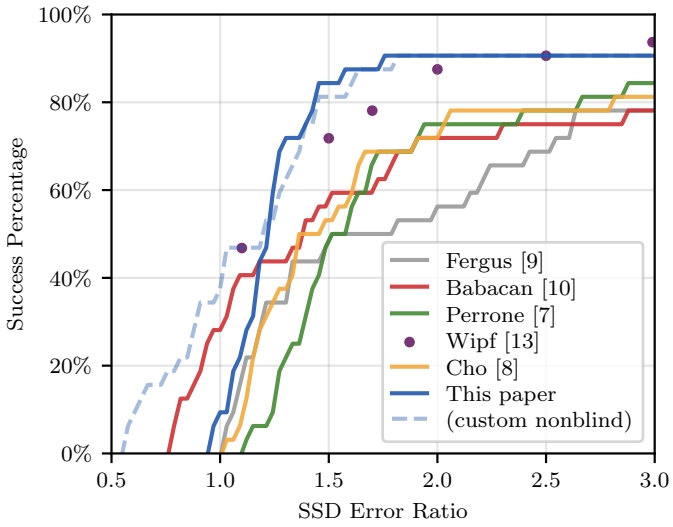


Fig. 3: Cumulative histogram of the error ratio on dataset [5].

Second, deblurring starts with a large value of σ_n^2 such as 10^{-3} (or even 10^{-2} in the downsampled problem) that exhibits the desired property $L(\mathbf{x}_{\mathbf{k}_0}, \mathbf{k}_0) < L(\mathbf{x}_\delta, \delta)$ and also leads to larger progress per IRCD iteration. Since a large σ_n^2 would lead to an oversmoothed image and a correspondingly inaccurate kernel estimate, σ_n^2 is then iteratively decreased during the optimization. This relies on the non-convexity of the cost function together with the monotonicity of IRCD [14] for staying in the desired basin of attraction.

We also add an L_2-penalty to the kernel as we have observed that this leads to somewhat "smoother" kernel estimates and slightly better overall results.[4]

In Fig. 3, we report results for the dataset [5] originally used to illustrate the inadequacy of naive MAP for blind image deblurring. All ratios between the sum-of-squared-distance (SSD) deblurring error of the estimated and true kernel were obtained using the non-blind deblurring algorithm of [17]—except for [10] who used their own non-blind deblurring algorithm. Our results appear to be at least on par with or better than any of the approaches discussed in Section II which attempt to avoid the MAP's perceived problems [9], [10], [12], [13].[5] This is in contrast to the naive MAP "proof-of-concept" results of [7], [8] which are significantly outperformed by [13].

Using their own non-blind deblurring stage allows [10] to deblur some of the images at an error ratio below one. To compare against their improved non-blind deblurring stage, Fig. 3 also includes results (dashed) of our method combined with a custom non-blind deblurring algorithm fixing $\mathbf{k}$ and using a variational representation of the Huber function [14], [18] for $\kappa(\cdot)$ in (4), which are superior to [10].

A visual example of our deblurring algorithm (using our own non-blind deblurring stage) is shown in Fig. 4. An

[4]We emphasize that this L_2-penalty is in no way necessary or essential for avoiding the no-blur solution in any of our experiments.

[5]The results of [12] are not depicted in Fig. 3 but are not superior to ours. Numerical results of [11] for the dataset [5] appear to be unavailable.

(a) Blurry image (b) Deblurred image and estimated kernel (c) Ground truth image and blur kernel

Fig. 4: Example from dataset [5].

additional example with higher resolution is given in Fig. 5, showing that naive MAP can also deblur larger images.

IV. Conclusion

We have revisited the original criticism [5] of the naive MAP approach for blind image deblurring and the subsequent literature. We have then attempted to show through conceptual analysis and experiments that this criticism is unwarranted. In evaluating our own variant of the naive MAP on the key dataset of [5] we have shown that a MAP approach not only avoids the no-blur solution but is able to produce high-quality results that are as good as any of the approaches that have been proposed as superior alternatives.

We do not claim that the naive MAP approach is superior in practice to other approaches, but we do claim that its reputation of being fundamentally flawed is unwarranted.

(a) Blurry image (b) Deblurred image (c) Ground truth

Fig. 5: A deblurring example with 453×600 pixels (best viewed on a computer screen).

References

[1] G. R. Ayers and J. C. Dainty, "Iterative blind deconvolution method and its applications," *Optics Letters*, vol. 13, no. 7, pp. 547–549, 1988.

[2] Q. Shan, J. Jia, and A. Agarwala, "High-quality motion deblurring from a single image," *ACM Trans. on Graphics*, vol. 27, no. 3, pp. 1–10, 2008.

[3] A. N. Rajagopalan and R. Chellappa, *Motion Deblurring: Algorithms and Systems*. Cambridge University Press, 2014.

[4] J. Huang and D. Mumford, "Statistics of natural images and models," in *Procs. IEEE Conf. on Computer Vision and Pattern Recognition*. IEEE, 1999, pp. 541–547.

[5] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman, "Understanding and evaluating blind deconvolution algorithms," in *Procs. IEEE Conf. on Computer Vision and Pattern Recognition*. IEEE, 2009, pp. 1964–1971.

[6] T. F. Chan and C.-K. Wong, "Total variation blind deconvolution," *IEEE Trans. on Image Processing*, vol. 7, no. 3, pp. 370–375, 1998.

[7] D. Perrone and P. Favaro, "Total variation blind deconvolution: the devil is in the details," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*. IEEE, 2014, pp. 2909–2916.

[8] S. Cho and S. Lee, "Convergence analysis of MAP based blur kernel estimation," in *Procs. IEEE Int. Conf. on Computer Vision*, 2017, pp. 4808–4816.

[9] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman, "Removing camera shake from a single photograph," in *Procs. ACM SIGGRAPH*, 2006, pp. 787–794.

[10] S. D. Babacan, R. Molina, M. N. Do, and A. K. Katsaggelos, "Bayesian blind deconvolution with general sparse image priors," in *Procs. European Conf. on Computer Vision*. Springer, 2012, pp. 341–355.

[11] L. Xu and J. Jia, "Two-phase kernel estimation for robust motion deblurring," in *Procs. European Conf. on Computer Vision*. Springer, 2010, pp. 157–170.

[12] L. Xu, S. Zheng, and J. Jia, "Unnatural L0 sparse representation for natural image deblurring," in *Procs. IEEE Conf. on Computer Vision and Pattern Recognition*, 2013, pp. 1107–1114.

[13] D. Wipf and H. Zhang, "Revisiting Bayesian blind deconvolution," *Journal of Machine Learning Research (JMLR)*, vol. 15, no. 111, pp. 3775–3814, 2014. [Online]. Available: http://jmlr.org/papers/v15/wipf14a.html

[14] H.-A. Loeliger, B. Ma, H. Malmberg, and F. Wadehn, "Factor graphs with NUV priors and iteratively reweighted descent for sparse least squares and more," in *Procs. IEEE Int. Symp. on Turbo Codes & Iterative Information Processing (ISTC)*. IEEE, 2018, pp. 1–5.

[15] J. Palmer, K. Kreutz-Delgado, B. D. Rao, and D. P. Wipf, "Variational EM algorithms for non-gaussian latent variable models," in *Adv. in Neural Information Processing Systems*, 2006, pp. 1059–1066.

[16] B. Ma, J. Trisovic, and H.-A. Loeliger, "Multi-image blind deblurring using a smoothed NUV prior and iteratively reweighted coordinate descent," in *Procs. IEEE Int. Conf. on Image Processing*. IEEE, 2020, pp. 973–977.

[17] A. Levin, Z. Weiss, F. Durand, and W. T. Freeman, "Efficient marginal likelihood optimization in blind deconvolution," in *Procs. IEEE Conf. on Computer Vision and Pattern Recognition*. IEEE, 2011, pp. 2657–2664.

[18] P. J. Huber and E. M. Ronchetti, *Robust Statistics*, 2nd ed. John Wiley & Sons, 2009.