Discriminative Learning of Iteration-wise Priors for Blind Deconvolution

Wangmeng Zuo[1], Dongwei Ren[1], Shuhang Gu[2], Liang Lin[3], Lei Zhang[2]
[1]School of Computer Science and Technology, Harbin Institute of Technology, Harbin, China
[2]Dept. of Computing, The Hong Kong Polytechnic University, Hong Kong, China
[3]Sun Yat-Sen University, Guangzhou, China

Blind deconvolution is a severely ill-conditioned problem, which under maximum a posterior (MAP) framework involves blur kernel estimation and non-blind deconvolution. For blur kernel estimation, sharp edge prediction and carefully designed image priors are vital to the success of MAP, guarding the solution free from the trivial delta kernel solution [3]. Edge prediction based approaches usually involve some heuristic and engineered methods, e.g., shock and bilateral filters [1], to restore salient edges explicitly. Image prior-based approaches generally deploy novel regularizers, e.g., l_1/l_2 [2], for implicit edge selection.

Recent researches also suggest that parameters for edge prediction and image regularization should be dynamically tuned during the iterations. For edge prediction-based approaches, strong edges are selected for coarse kernel estimation in the first a few iterations, and subsequently more details are added to further refine the estimated kernel [1]. Under the MAP framework, the regularization parameter λ is set small to preserve strong edges while suppressing detailed textures in the first a few iterations, and gradually tuned along with the iteration to produce accurate kernel [2].

Rather than the fixed regulariers used in previous work, in this paper we propose iteration-wise priors to adaptively select appropriate edges to facilitate estimate kernel $\mathbf{k}$. By modeling the image gradients $\mathbf{d}$ as Laplacian distribution $\Pr(\mathbf{d}) \propto e^{-\|\mathbf{d}\|_p^p/\lambda}$, the proposed model is formulated as

$$\begin{array}{l} \min_{\mathbf{d},\mathbf{k}} \frac{\lambda^{(t)}}{2\sigma_n^2} \|\mathbf{k}\otimes\mathbf{d}-\nabla\mathbf{y}\|^2 + \|\mathbf{d}\|_{p^{(t)}}^{p^{(t)}} + \mu \|\mathbf{k}\|_{0.5}^{0.5} \\ \text{s.t. } \nabla_h \mathbf{d}_v = \nabla_v \mathbf{d}_h, \sum_i k_i = 1, k_i \geq 0, \forall i. \end{array} \quad (1)$$

where the parameters $\{\lambda^{(t)}, p^{(t)}\}$ for iteration t is uniquely set.

To avoid heavy handcrafted parameter tuning, we propose a principled discriminative approach to learn the parameters from synthetic training dataset. Denote $\mathcal{D}$ by a set of synthetic images $\{(\mathbf{d}_i^{gt}, \mathbf{k}_i^{gt}, \nabla\mathbf{y}_i)\}_{i=1}^N$, where $\mathbf{d}_i^{gt}$ denotes the gradient of the i-th clear image, $\mathbf{k}_i^{gt}$ denotes the i-th blur kernel, and $\nabla\mathbf{y}_i$ denotes the gradient of the i-th blurry image. For estimating $\boldsymbol{\theta}^{(t)} = \left\{\lambda^{(t)}, p^{(t)}\right\}$, we adopt the weighted mean square error (MSE) loss function defined on $\mathcal{D}$,

$$\begin{array}{l} L^{(t)}(\boldsymbol{\theta}) = \sum_{i=1}^N L_i^{(t)}(\boldsymbol{\theta}) = \sum_{i=1}^N \alpha^{(t)} L_{\mathbf{d}_i}^{(t)}(\boldsymbol{\theta}) + L_{\mathbf{k}_i}^{(t)}(\boldsymbol{\theta}) \\ = \sum_{i=1}^N \frac{\alpha^{(t)}}{2} \left\| \mathbf{d}_i^{(t)} - \mathbf{d}_i^{gt} \right\|^2 / \left| \mathbf{d}_i^{gt} \right| + \frac{1}{2} \left\| \mathbf{k}_i^{(t)} - \mathbf{k}_i^{gt} \right\|^2 / \left| \mathbf{k}_i^{gt} \right|, \end{array} \quad (2)$$

where $|\bullet|$ counts the entries of the vector for the normalization of image and kernel sizes, and α denotes the trade-off parameter. Due to the specially designed one-step augmented Lagrangian solutions to kernel estimation and non-blind deconvolution, the optimal parameters for each iteration can be searched via simple gradient descent method, Interestingly, thanks to the generalized shrinkage / threshold (GST) operator [7] for non-convex l_p optimization, the case $p < 0$ is allowed, which magnified the salient edges while suppressing harmful detailed textures, so that the coarse shape of blur kernel can be rapidly estimated.

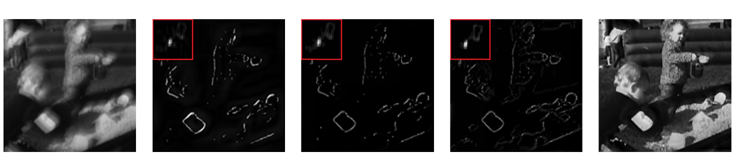

Figure 1: Illustration of intermediate kernel estimation with learned p values.

This is an extended abstract. The full paper is available at the Computer Vision Foundation webpage.

The iteration-wise parameters are learned from Levin et al.'s dataset [3]. As shown in Fig. 1, we observe that the small λ and p values benefit the rough kernel estimation, and with the values increasing the estimated kernel is refined in details. Also at beginning, $p < 0$ magnified the preserved salient edges, significantly contributing to rapid kernel estimation. The learned parameters can be directly applied to other synthetic and real blurry images. In our experiments, the quantitative evaluation is conducted on Levin et al.'s dataset [3] and Sun et al.'s dataset [5]. From Table 1 and 2, the proposed method achieves better quantitative metrics than the existing gradient prior-based methods, including both MAP [1, 2, 6] and variational Bayes [4], and is comparable with the state-of-the-art patch-based method [5], but is much more efficient. We also provide 2 variants of our method with p fixed as -1 and 0.2 on Sun et al.'s dataset, via which the superiority of iteration-wise priors is validated as shown in Table 2. More deblurring results of synthetic and real blurry images are provided in our manuscript and supplementary, and the deblurring results by the proposed method are more visually plausible.

Table 1: Comparisons on Levin et al.'s dataset [3] using mean PSNR, mean SSIM, mean error ratio and mean running time (seconds)

	PSNR	SSIM	Error Ratio	Time
Known **k**	32.31	0.9385	1.0000	—
Krishnan et al. [2]	28.26	0.8547	2.3746	8.9400
Cho & Lee [1]	28.83	0.8801	1.5402	1.3951
Levin et al. [4]	28.79	0.8922	1.5592	78.263
Xu & Jia [6]	29.45	0.9000	1.4071	1.1840
Sun et al. [5]	30.85	0.9191	1.2244	191.03
Ours	30.33	0.9192	1.2537	25.184

Table 2: Comparisons on Sun et al.'s dataset [5] using mean PSNR, mean SSIM, mean error ratio and mean running time (seconds)

	PSNR	SSIM	Error Ratio	Time
Known **k**	32.35	0.9536	1.0000	—
Krishnan et al. [2]	22.76	0.8136	6.8351	159.29
Cho & Lee [1]	26.13	0.8624	5.0731	10.518
Levin et al. [4]	24.64	0.8606	4.5798	518.59
Xu & Jia [6]	28.11	0.9016	3.2843	6.2940
Sun et al. [5]	29.32	0.9200	2.4036	3911.1
Ours (-1)	27.96	0.9019	3.2188	311.77
Ours (0.2)	28.35	0.9111	2.9877	312.11
Ours	29.10	0.9220	2.4054	311.61

[1] S. Cho and S. Lee. Fast motion deblurring. *ACM Transactions on Graphics (TOG)*, 28(5):145, 2009.
[2] D. Krishnan, T. Tay, and R. Fergus. Blind deconvolution using a normalized sparsity measure. In *CVPR*, 2011.
[3] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman. Understanding and evaluating blind deconvolution algorithms. In *CVPR*, 2009.
[4] Anat Levin, Yair Weiss, Fredo Durand, and William T Freeman. Efficient marginal likelihood optimization in blind deconvolution. In *CVPR*, 2011.
[5] Libin Sun, Sunghyun Cho, Jue Wang, and James Hays. Edge-based blur kernel estimation using patch priors. In *ICCP*, 2013.
[6] L. Xu and J. Jia. Two-phase kernel estimation for robust motion deblurring. In *ECCV*, 2010.
[7] Wangmeng Zuo, Deyu Meng, Lei Zhang, Xiangchu Feng, and David Zhang. A generalized iterated shrinkage algorithm for non-convex sparse coding. In *ICCV*, 2013.