Adaptive Shrinkage Cascades for Blind Image Deconvolution

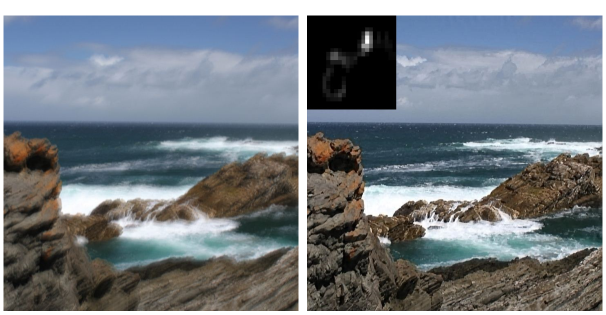
Xuejian Rong and Yingli Tian
Department of Electrical Engineering
The City College, City University of New York, USA
{xrong, ytian}@ccny.cuny.edu

***Abstract*—Recently emerged discriminative non-blind deconvolution methods achieve excellent performance with only a fraction of computation cost w.r.t. generative competitors, but their extension to blind deconvolution field was seldom addressed in a practical manner, albeit equally vital in image restoration area. We propose a novel framework for effective blind image deblurring by patch-wise prior based adaptive shrinkage cascades, which introduces the powerful internal patch-based image statistics to the non-blind shrinkage field formulations. The rich expressiveness of internal patch prior brings instance-specific adaptivity to alternating kernel refinement between neighboring shrinkage cascades, while shrinkage model trained from varieties of natural image collections benefits internal patch-wise prior inference with external information and superior efficiency.**

***Index Terms*—Blind Image Deblurring; Image Deconvolution; Image Restoration;**

I. Introduction

Camera shake during the exposure time in digital photography often results in undesirable image degradations. Although showing effectiveness on a certain extent, current stabilization techniques embedded with modern cameras and smart devices still cannot handle relatively large and/or rapid camera motion (see Figure 1a). Recapturing is also usually impossible after realizing the blurry results. Blind image deconvolution aims to solve this problem by estimating a latent sharp version of the blurry input image when the blur kernel is unknown, which is a severely ill-posed problem due to the ambiguity. Even when the ideal blur kernel was known beforehand (*i.e.*, non-blind deconvolution), directly doing the inversion will still yield an unpleasing result lacking high-frequency details, due to the corruption of inevitably existing unknown noise and outliers.



(a) blurry image (b) deblurred image

Fig. 1: Blind deconvolution results through our adaptive shrinkage cascades framework. The predicted blur kernel is shown at the top-left corner of (b).

To solve this ambiguity problem, varieties of prior knowledge, namely, regularizer, have been proposed to refine the restoration process in last decades. Currently, most successful blind deconvolution approaches aim to estimate the blur kernel (*i.e.*, point spread function) first [2], [8], [16], [22], which is then fixed to bootstrap the non-blind deconvolution process. On the one hand, patch-wise approaches, e.g., [18], [24], attract even more interests than pixel-wise competitors due to their rich expressiveness. On the other hand, despite existing *generative* non-blind deblurring approaches, such as TV-L1 model [23] and Gaussian Markov random fields [14], achieve good deblurring results, their complicated optimization scheme restricts their extension to the real world applications. In contrast, newly emerged *discriminative* non-blind deconvolution methods, including the neural network-based scheme [15], [17], regression tree fields (RTF)-based prediction cascades [13], and shrinkage fields [12], achieve impressive image restoration quality comparable with, or surpassing, the current state-of-the-art generative competitors. However, considering the increasing importance of discriminative approaches, their extensions to united blind image deconvolution are still relatively ignored, especially for the relationship with the patch-wise blur estimation schemes.

In this paper, we propose an adaptive shrinkage cascade approach for blind image deconvolution inspired by the non-blind shrinkage field deconvolution framework [12]. Our proposed approach can adapt to specific deblurring target due to the internal image patch recurrence property, while keeping the superior quality of shrinkage fields. We also design an efficient optimization framework to ensure the minimization convergence. In particular, Schelten *et al.* [11] first extended the non-blind discriminative RTF framework to blind deblurring tasks by interleaving the RTF cascades with standard kernel updating, but the interleaved RTF cascades still suffer from complicated field structure, and the trained model lacks the instance-specific adaptivity as our proposed framework. Wang *et al.* [21] attempted to combine the power of discriminative random fields with patch-based deconvolution methods, by using RTF model to index the Expected Patch Log-Likelihood (EPLL) based patch priors [24]. However, the refined model still relies on the complex Gaussian mixture model (GMM) learning the scheme, and lacks inter-related blur kernel estimations.

The proposed framework generalizes the discriminative non-

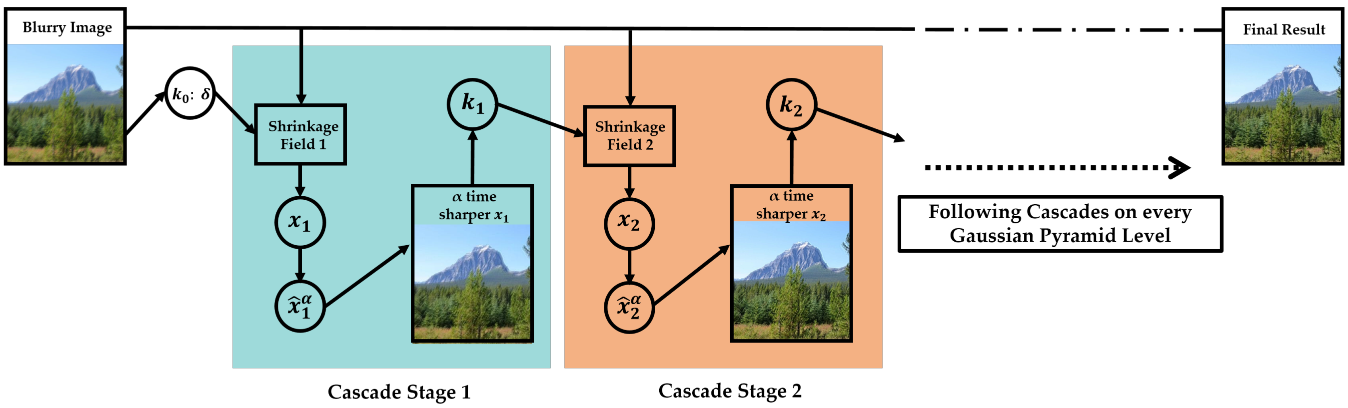


Fig. 2: Flowchart of the proposed adaptive shrinkage cascade framework. The input blurry image x is shrunk by a factor of α, then the result x^* contains α-times less the amount of blur. The final deblurred result is generated through all the cascade stages in a multi-scale manner.

blind deconvolution procedure to the blind deconvolution with the patch-based blur kernel inference. It also reveals more insights on the connection between discriminative and patch-wise optimization methods. Our main contributions include:

- A novel framework which bridges the gap between patch-based prior inference scheme and discriminative non-blind deconvolution methods;
- An effective way of constructing alternating discriminative deconvolution scheme which leverages the internal image patch recurrence statistics.

II. Related Work

In the blur kernel estimation field, the pixel-wise image prior has been extensively studied, including the heavy-tailed gradient distribution [2], hyper-Laplacian distributions [7], salient edge prior [22], or the natural image power law [3]. Recently, patch-wise image prior has attracted particular attention since its larger spatial support could model complex image structures and dependencies in larger neighborhoods, thus faithfully describe the prior knowledge in middle and high levels. Sun *et al.* [18], [19] further extended this framework by augmenting the single-scale patch priors to a multi-scale formulation. Without relying on a large scale external image training dataset, Michaeli *et al.* [9] focused on studying deblurring using the internal image statistics such as the internal patch recurrence property. This method better copes with the self-similarity structures including the relatively smooth scenes of sky, ocean or square, which cannot be handled by external patch-based competitors.

On the deconvolution stage, conditional random fields (CRFs) are often employed in discriminative image restoration methods. Tappen *et al.* [20] firstly proposed highly efficient discriminatively trained Gaussian CRFs. Recently the regression tree fields (RTFs) [6] methods are widely used, in which the parameters of these Gaussian CRFs are determined by non-parametric regression trees. Its effectiveness and efficiency have been successfully validated in varieties of restoration tasks, especially for deconvolution and denoising [5], [6]. Recently, existing discriminative non-blind deblurring approaches rely either on neural network [15], [17], or conditional random field (CRF) cascades [12], [13]. We hereby propose a new method based on shrinkage field cascades [12], since the stacked structure could effectively benefit the blind kernel refinement process.

III. Shrinkage Cascade for Blind Deconvolution

A. Non-blind Shrinkage Fields

To effectively extend the discriminative non-blind deconvolution procedure to blind deconvolution field, we first analyze the structure of shrinkage fields, and how it could be adapted into a stage refinement manner with patch prior based kernel estimation, instead of naive direct concatenation. First we model the camera shake in a blurry image by employing a conventional formation which models image blur as convolution under additive noise: $y = x \otimes k + \varepsilon$, where x denotes the latent image, k is the blur kernel, ε as the image noise, $\otimes$ for the convolution process and y is the observed blurry image. The latent sharp image is then restored x from its corrupted observation y by combining an observation likelihood and an image prior: $p(x|y) \propto p(y|x) \cdot p(x)$, here the corruption process is modeled with a Gaussian likelihood. To minimize the posterior distribution $p(x|y) \propto \exp(-E(x|y))$, directly employing gradient-descent algorithms is not satisfying enough. We hereby rely on the half-quadratic optimization which introduces independent auxiliary variables z_{ic} for all filter responses $\eta = f_i^T x_c$ over all cliques c to obtain an augmented energy function $E(x, z|y)$.

$$E(x, z|y) = \frac{\lambda}{2}\|y - Kx\|^2 + \sum_{i=1}^{N}\sum_{c \in C}\left[\frac{\beta}{2}(\eta - z_{ic})^2 + \rho_i(z_{ic})\right], \quad (1)$$

where the parameters β and ρ are introduced as the constraints. z_{ic} denotes the auxiliary variables and η affects the corresponding filter responses. K denotes the convolution matrices corresponding to the blur kernel k, and λ is related to the strength of the assumed additive Gaussian noise. Regularization is provided through a Markov random field model (fields of experts [10]) with robust potential functions ρ_i. The further alternating minimization procedure replaces the conventional shrinkage function in the wavelet image restoration literature [4] with a flexible shrinkage function modeled as a linear combination of Gaussian RBF kernels, and at the same time keeps the purpose to shrink small filter/wavelet coefficients by pulling them towards zero, because they are assumed to be caused by noise instead of signal.

Since half-quadratic optimization typically involves several iterations, multiple predictions could be chained as a cascade of shrinkage fields. Therefore, we aim to interleave the patch-wise kernel refinement procedure between different cascade, to further effectively boost the performance of overall deconvolution system, which is described in details in Section III-C.

B. Internal Patch Recurrence of Natural images

Due to the impressive performance and efficiency of the discriminative shrinkage field non-blind deconvolution framework, combining their power with compatible blur kernel estimation scheme seems to be a natural idea. Directly input the blur kernel estimated by classical leading pixel-wise blur estimation approaches into the discriminative CRF scheme does not perform well [12], since the CRF framework associates nonadjacent pixels by connecting them in the field and needs larger spatial support. For the patch-wise schemes, the internal patch-based method [9] only depends on the internal statistics of the blurry image itself, therefore, cannot benefit from the externally learned information to restore detailed high-frequency textures, thus compromises image details for smoothness. This could be compensated by the discriminative shrinkage field model learned from a dataset with rich varieties of natural images.

Internal patch-based deblurring methods aim to take full advantage of the cross-scale recurrence property which most natural images obey but diminish in blurry images. Thus, the deviations from ideal patch recurrence could be exploited as a clue for recovering the underlying unknown blur kernel. In particular, while the blur is strong in the original input blurry image, the blur decreases at coarser scales of the image. Specifically, if we shrink a blurry image x by a factor of α, then the result x^* contains α-times less the amount of blur, as illustrated in Figure 2. These patches in coarser image scales can thus serve as a good patch prior for deblurring the input scales.

As the internal patch pool constructed based on every target input blurry image is context-appropriate and self-adaptive, the internal patch-based method handles the self-similarity blurry regions and other instance-specific image structures well, which could compensate the generally trained discriminative shrinkage field deconvolution model by further tuning the shrinkage stages in the model at the testing stage (see Section III-C).

C. Adaptive Shrinkage Cascades

Although the internal patch-based blur estimation is able to provide decent kernel predictions for the discriminative shrinkage field deconvolution framework, a naive combination of the two estimation and deconvolution methods methods will not lead to a satisfactory improvement in the final deblurring results, because the blur kernel estimated beforehand is fixed in the shrinkage field deconvolution process, like the RTF cascades in [13]. In this case, re-estimation and refinement of the blur kernel during different cascades of shrinkage fields is necessary and beneficial. In this case, we tend to embed the internal patch-wise kernel refinement between neighboring shrinkage cascades, and tune the pre-trained shrinkage cascade model with the instance-specific adaptation property of internal patch recurrence statistics.

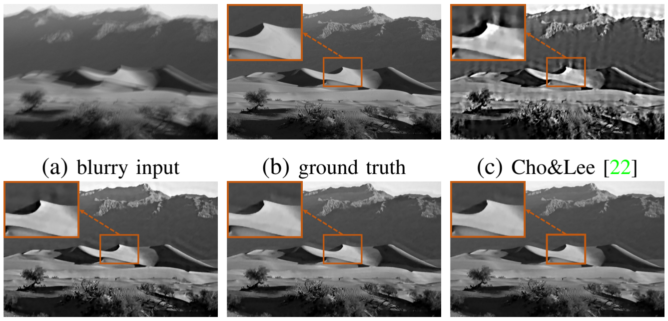
More specifically, we seek an image $\widehat{x}$ and a blur kernel $\widehat{k}$ such that $\widehat{x}$ satisfies the patch recurrence property, (i.e., strong similarity between patches across scales of $\widehat{x}$) as well as $\widehat{k}\otimes\widehat{x}$ approaches the blurry image y as much as possible. At every shrinkage cascade stage, we compute the kernel update using the cross-scale patch recurrence refinement scheme by minimizing the following objective function with respect to k:

$$\arg\min_{\widehat{x},k} \|y - k \otimes \widehat{x}\|^2 + \lambda_1 \rho(\widehat{x}, \widehat{x}^{\alpha}) + \lambda_2 \|k\|^2 . \quad (2)$$

Hereby, the $\widehat{x}^{\alpha}$ is an α-times smaller version of $\widehat{x}$. The second term $p(x, x^{\alpha})$ measures the degree of dissimilarity defined as the minus EPLL between patches in x and their nearest neighbor patches in x_{α}. The third term is a regularizer on the kernel k.

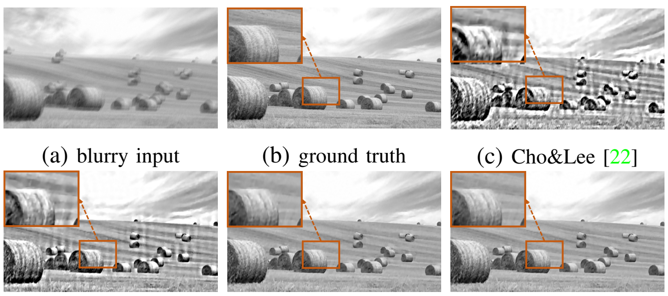
$$\begin{aligned} \rho(\widehat{x}, \widehat{x}^{\alpha}) &= -\sum_j \log\left(\sum_i \exp\{\Gamma\}\right), \\ \Gamma &= -\frac{1}{2h^2} \|Q_j \widehat{x} - R_i \widehat{x}^a\|^2 , \end{aligned} \quad (3)$$

where h is a bandwidth parameter, Q_j and R_i are matrices that extract the corresponding patches from $\widehat{x}$ and $\widehat{x^{\alpha}}$ respectively. Here we initialize the whole blind adaptive shrinkage cascade framework with a delta function kernel δ, then refine the kernel over the shrinkage cascades in a coarse-to-fine scales of image pyramid scheme, i.e., the final kernel estimate at one scale is upsampled to serve as the initial estimate of the next scale of shrinkage cascades. This alternating minimization procedure is essential for updating the $\rho(x, x^{\alpha})$, because the EPLL metric makes the objective function non-convex, thus no closed-form solution could be directly gotten. In the other hand, EPLL introduces the patch-based probabilistic prior to our framework. Figure 2 illustrates the whole scheme of the adaptive shrinkage cascades. Hereby, shrinkage cascades are used to predict image and kernel estimates at each level of a Gaussian pyramid. The estimates at one level are enlarged to serve as inputs for the next finer level. Going from coarse to fine, we train progressively more powerful shrinkage models to account for the higher level of image details.



(a) blurry input (b) ground truth (c) Cho&Lee [22]
(d) Xu&Jia [23] (e) Michael [9] (f) Proposed

Fig. 3: Comparison on uniformly blurry synthetic natural image in [18].



(a) blurry input (b) ground truth (c) Cho&Lee [22]
(d) Xu&Jia [23] (e) Michaeli [9] (f) Proposed

Fig. 4: Comparison on synthetic natural scene image similar to Figure 3.

D. Learning the Shrinkage Cascade Parameters

The adaptive shrinkage cascades excel due to their adaptability to kernel estimation, and yield better results when trained with rich varieties of natural image collections. Here we train the cascade model based on the Berkeley Segmentation Dataset [1], by downsampling all 500 grayscale images to half size, compressing outliers, and reducing noise, then getting the synthetic blurry images with selected kernels which are totally different from the ones in testing dataset.

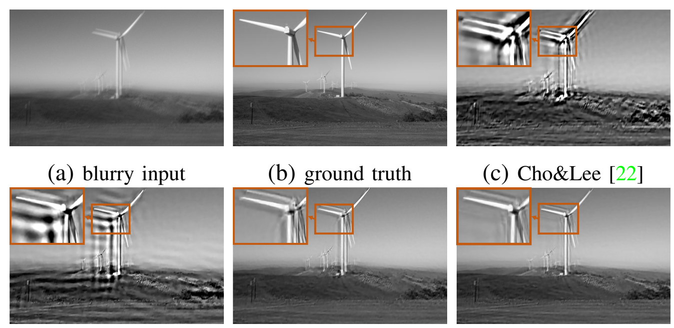
To leverage more discriminative features than simple pixel intensities, we adopt the fields of experts (FoE) filter bank of [10]. By learning different model parameters for every stage of our cascade, we essentially learn random field models for each iteration of the corresponding optimization algorithm. As [23] suggested, we train our model with a mixture of perfect and estimated kernels instead of the ground truth kernel, since it helps to adapt to kernel estimation errors at the test stage and achieves superior performance.

In the non-blind deconvolution part, we follow [13] and parameterize the prediction with the blur kernel, such that the instance-specific blur is provided at inference. Although our model is not trained for any specific blurs, it could still benefit from the context-aware internal patch recurrence property as indicated in Section III-B. Here the shrinkage functions are differentiable and do not need time-consuming custom optimization like [11].

IV. Experiments

In this section, the proposed framework is evaluated on multiple synthetic blurry images, including natural and man-made scenes. To guarantee the fair comparison, we bootstrap our scheme with the rough kernel from [23] as Schmidt *et al.* [12] and Schelten *et al.* [11] do. Figures 3 and 4 demonstrate that our method preserves challenging regions of image texture, while suppressing ringing and noise artifacts in smooth regions or on the image boundary, which shows visibly superior performance of the shrinkage cascades over a wide variety of blind deconvolution methods.

Figure 5 illustrates that the instance specific adaptivity in our framework can deal with relatively smooth regions in blurry images. While other methods stuck in the difficult self-similar sky region and produce unsatisfying results, our proposed adaptive shrinkage cascades substantially achieve better deblurring results due to the adaptivity to varying structures, especially for self-similarity. The zoomed-in area shows that Michaeli *et al.* [9]'s framework also handles the smooth region well, but the results seem to be slightly over-smoothed and lack details.



(a) blurry input (b) ground truth (c) Cho&Lee [22]
(d) Xu&Jia [23] (e) Michaeli [9] (f) Proposed

Fig. 5: Comparison on an blurry image in [18] with explicits self-similar regions, which demonstrates the effectiveness of the proposed framework on varying structures and relatively smooth regions in blurry images.

V. Conclusion

In this paper, we have investigated the underlying connection between the patch-based prior inference scheme and discriminative non-blind deconvolution method, and have constructed an effective alternating deconvolution scheme which takes full advantage of the internal image patch recurrence property. The experiment results visually validate the effectiveness of the proposed adaptive shrinkage cascade framework. It will also help improve the performance by introducing the power of Convolutional Neural Networks (CNN) for the blur kernel prediction in the future.

VI. Acknowledgement

This work was supported in part by NSF grants EFRI-1137172, IIP-1343402, and FHWA grant DTFH61-12-H-00002.

REFERENCES

[1] P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik. Contour detection and hierarchical image segmentation. *IEEE TPAMI*, 33(5):898–916, May 2011.

[2] R. Fergus, B. Singh, A. Hertzmann, S. Roweis, and W. T. Freeman. Removing camera shake from a single photograph. In *ECCV 2014*.

[3] A. Goldstein and R. Fattal. Blur-kernel estimation from spectral irregularities. In *ECCV 2012*.

[4] Y. Hel-Or and D. Shaked. A discriminative approach for wavelet denoising. *IEEE TIP*, 17(4):443–457, 2008.

[5] S. N. J. Jancsary and C. Rother. Loss-specific training of non-parametric image restoration models: A new state of the art. In *ECCV 2012*.

[6] J. Jancsary, S. Nowozin, T. Sharp, and C. Rother. Regression tree fields – an efficient, non-parametric approach to image labeling problems. In *CVPR 2012*.

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

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

[9] T. Michaeli and M. Irani. Blind deblurring using internal patch recurrence. In *ECCV 2014*.

[10] S. Roth and M. J. Black. Fields of experts. *IJCV*, 82(2):205–229, Apr. 2009.

[11] K. Schelten, S. Nowozin, J. Jancsary, C. Rother, and S. Roth. Interleaved regression tree field cascades for blind image deconvolution. In *WACV 2015*.

[12] U. Schmidt and S. Roth. Shrinkage fields for effective image restoration. In *CVPR 2014*.

[13] U. Schmidt, C. Rother, S. Nowozin, J. Jancsary, and S. Roth. Discriminative non-blind deblurring. In *CVPR 2013*.

[14] U. Schmidt, K. Schelten, and S. Roth. Bayesian deblurring with integrated noise estimation. In *CVPR 2011*.

[15] C. Schuler, H. C. Burger, S. Harmeling, and B. Scholkopf. A machine learning approach for non-blind image deconvolution. In *CVPR 2013*.

[16] Q. Shan, J. Jia, and A. Agarwala. High-quality motion deblurring from a single image. *ACM T. Graphics*, 27(3), 2008.

[17] J. Sun, W. Cao, Z. Xu, and J. Ponce. Learning a Convolutional Neural Network for Non-uniform Motion Blur Removal. In *CVPR 2015*.

[18] L. Sun, S. Cho, J. Wang, and J. Hays. Edge-based blur kernel estimation using patch priors. In *ICCP 2013*.

[19] L. Sun, S. Cho, J. Wang, and J. Hays. Good image priors for non-blind deconvoluton: Generic vs specific. In *ECCV 2014*.

[20] M. F. Tappen, C. Liu, E. H. Adelson, , and W. T. Freeman. Blur-kernel estimation from spectral irregularities. In *CVPR 2007*.

[21] Y. Wang, S. Cho, J. Wang, and S. Chang. Discriminative indexing for probabilistic image patch priors. In *ECCV 2014*.

[22] Y. Wang, S. Cho, J. Wang, and S.-F. Chang. Discriminative indexing for probabilistic image patch priors. In *ECCV 2014*.

[23] L. Xu and J. Jia. Two-phase kernel estimation for robust motion deblurring. In *ECCV 2010*.

[24] D. Zoran and Y. Weiss. From learning models of natural image patches to whole image restoration. In *ICCV 2011*.