# Text Image Deblurring

Wang Yuzhen ID:2018E8018461008

***Abstract*— Camera shake during exposure time often results in spatially variant blur effect of the image. There are many blind natural image deconvolution approaches. They can't deal with blurry text images well, since they rely on natural image statistics which do not respect the properties of text images. On the other hand, these methods can't handle the large motion blurs and complex background. We use a simple effective $L_0$-regularized prior based on intensity and gradient for text image deblurring. We discuss the relationship with edge-based deblurring methods and present how to select salient edges more principally. Experimental results demonstrate that the proposed algorithm performs well.**

***Index Terms*— Blurry kernel MAP $L_0$-regularized prior image statistics**

## I. Introduction

THE motion deblurring problem has drawn much attention in recent years with demonstrated success. Most approaches model a blurry natural image as the integration of intermediate frames captured by the camera along the motion trajectory, typically assuming in-plane translational camera motions. Liked the Figure 1, it is a simple image formation process with convolution and blur kernel. Since blind deconvolution is an ill-posed problem, prior knowledge or additional information are often required for effective solutions.

There are some early methods using Richardson–Lucy deconvolution (RL) [1], [2] or Wiener filter [3]. Early work splits an image into several regions where each is modeled with a constant PSF and recovered by a uniform deblur algorithm. Then, the great success of the Maximum-a-Posteriori (MAP) algorithms [4], [5], [6], [7], [8], [9], [10], can be attributed to the use of statistical priors on natural images and selection of salient edges for kernel estimation. About the MAP, a good method is from image prior. We will talk in the section 2.

If a motion blur is shift-invariant, it can be modeled as the convolution of a latent image with a motion blur kernel, where the kernel describes the trace of a sensor in Fig. 1. Then, removing a motion blur from an image becomes a deconvolution operation. In non-blind deconvolution, the motion blur kernel is given and the problem is to recover the latent image from a blurry version using the kernel.

In blind deconvolution, the kernel is unknown and the recovery of the latent image becomes more challenging. In this paper, we solve the blind deconvolution problem of a single image, where both blur kernel and latent image are estimated from an input blurred image.

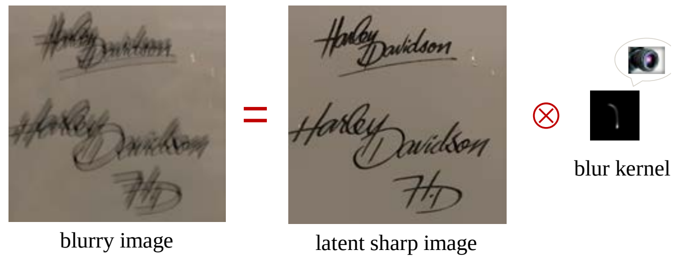


Fig. 1. Image formation process

The problem that general image deblurring approaches face when dealing with text images is the way they handle small objects. Since, text images are different from natural images.

Previous document restoration methods [11, 12, 13, 14] have used the two-tone or bi-level property of text. Under this assumption, small blurs and ink-bleeding artifacts can be removed using thresholding. However, such thresholding no longer works when the image is blurred with a large motion kernel. But the gradient histogram of a natural image is significantly different from a text image. Furthermore, strong gradients of a text image have a highly-regular spatial distribution as a prior. We will define the properties text image.

## II. Image Prior

In order to estimate the latent image from such limited measurements, it is essential to have some notion of which images are a priori more likely.

### A. Natural image statistics

For example, Figure 2 shows a natural image and a histogram of its gradient magnitudes. The distribution shows that the image contains primarily small or zero gradients, but a few gradients have large magnitudes. Fergus represents the distribution over gradient magnitudes with a zero mean mixture-of-Gaussians model, as illustrated in Figure 2. This representation was chosen because it can provide a good approximation to the empirical distribution. The prior can be used for deblurring natural images well.

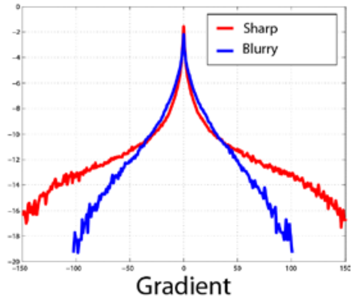


Fig. 2. Natural images characteristic distribution with heavy tails

### B. Text image statistics

Because natural image statistics does not work well for text image deblurring, we read papers about text-specific properties that should be considered in the deblurring process.

In Cho et al. [15], text-specific properties are listed.

**Property 1.** Text characters usually have high contrasts against nearby background regions.

**Property 2.** Each character has a near-uniform color. Gradient values inside each character should be close to zero.

**Property 3.** Although many documents have single background colors (e.g., white), advertisements or posters may have more complex background. Thus, assuming a single tone for background as in [1] is too restrictive. We instead assume the background gradient values obey natural image statistics and are sparse.

### *C.* L0-Regularized Intensity and Gradient Prior

In Pan et al. [15], the proposed intensity and gradient prior is based on the observation that text and background regions usually have nearly uniform intensity values in clear images without blurs. Fig. 2b illustrates that the pixel intensity distribution of a clear text image (Fig. 2a) is peaked at two values (near 0 and 255). For a blurred text image, the pixel intensity distribution is significantly different from that of a clear image. Fig. 2e shows the histogram of pixel intensities (from a blurred image in (d)) with fewer pixels of value 0 and 255. The reason is that each pixel in a blurred image can be viewed as the weighted sum of a few neighborhood pixels of a clear image.

Thus, the intensity distribution is squeezed from both ends of the intensity range. As a result, there are fewer pure black pixels (intensity value 0) in a blurred text image than a clear one. It is clear that the nonzero values of blurred image gradients are denser than those of the clear one. These properties do hold for generic images, and appear more obviously in text images.

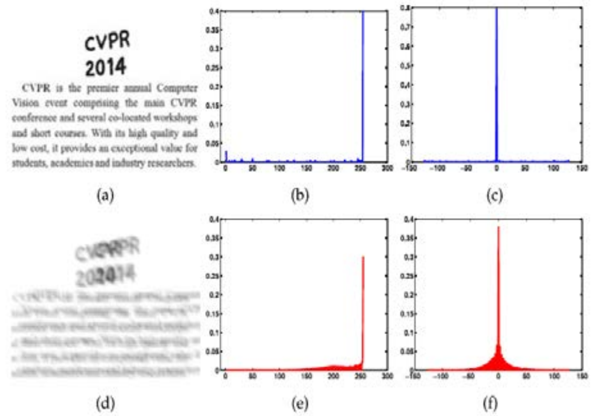


Fig. 3. Pan et al. Intensity and gradient properties of text images. (a) A clear text image. (b) Pixel intensity distribution from (a). (c) Distribution of horizontal gradient from (a). (d) A blurred image. (e) Pixel intensity distribution from (d). (f) Distribution of horizontal gradient from (d).

- L0-norm:

$$\rho(x) = \|x\|^0 = \begin{cases} 0, & |x| = 0 \\ 1, & |x| \neq 0 \end{cases}$$

- Approximate L0 sparsity function:

$$\phi(\nabla_* I) = \begin{cases} \frac{1}{\epsilon^2}|\nabla_* I|^2, & if\ \ |\nabla_* I| \le \epsilon \\ 1, & otherwise \end{cases}$$

## III. Model and Method

Our algorithm takes as input a blurred input image B, which is assumed to have been generated by convolution of a blur kernel K with a latent image L plus noise:

$$B = L \otimes K + N \quad (1)$$

where B is a blurred image, K is a motion blur kernel or a point spread function, L is a latent image, N is unknown noise introduced during image acquisition, and ⊗ is the convolution operator. In blind deconvolution, we estimate both K and L from B, which is a severely ill-posed problem.

Main objective function:

$$(L^*,\ K^*) = arg\min_{L,K}\ \|B - L \otimes K\|_2^2 + \ \rho_L(L) + \rho_K(K) \quad (2)$$

$\rho_L(L)$ : prior of latent image

$\rho_K(K)$ : prior of kernel

In the latent image estimation and kernel estimation steps of the process, we respectively solve the equations similar to a non-linear problem, iterative optimization :

$$L^* = arg\min_{L}\ \|B - L \otimes K\|_2^2 + \ \rho_L(L) \quad (3)$$

$$K^* = arg\min_{K}\ \|B - L \otimes K\|_2^2 + \ \rho_K(K) \quad (4)$$

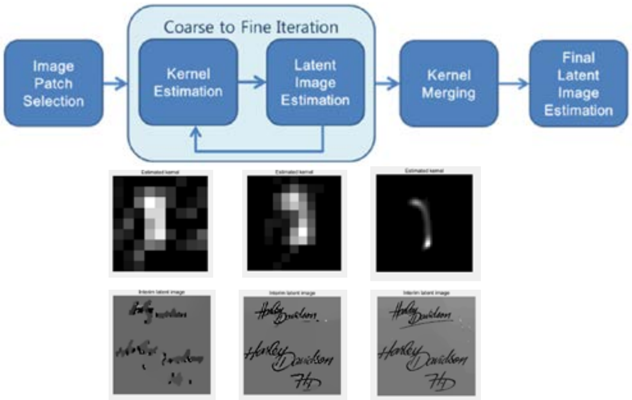


Fig. 4. A new optimization framework to incorporate the text properties in deblurring

## *A.* Estimating L

Restore sharp edges along the boundaries of texts considering the text properties 1 and 2. Due to the L0 regularization term, minimizing (2) is computationally intractable. Based on the half-quadratic splitting L0 minimization approach [33], we propose an efficient alternating minimization method to solve this problem.

The objective function can be rewritten as

$$I^{(t+1)} = arg\min_{I}\left\{\left\|K^{(t)} \otimes I - B\right\|_2^2 + \lambda \sum_{*\in\{x,y\}} \phi_0(\partial_* I)\right\}.$$

## *B.* Estimating K

Compute the blur kernel (camera's trajectory) . Given x, (6) is a least squares minimization problem in which a closed-form solution can be computed by FFTs. As the estimation based on gradients has been shown to be more accurate [3], [5], [7], we estimate the blur kernel k by

$$K^{(t+1)} = arg\min_{K}\left\{ \left\|K \otimes I^{(t+1)} - B\right\|_2^2 + \gamma\|K\|_2^2 \right\}$$

## C. Final deconvolution

$$E(L) = \sum_{\partial^*} \|\partial^* B - \partial^* L \otimes K\|_2^2 + \lambda\, \|\nabla L\|_2^2$$

## D. main processes

Fig. 4 shows the overall process of our blind deconvolution method. An algorithm summarizing the process can also be found in the supplementary material. To progressively refine the motion blur kernel K and the latent image L, our method iterates three steps: prediction, kernel estimation, and deconvolution. Recall that our latent image estimation is divided into prediction and deconvolution. We place prediction at the beginning of the loop to provide an initial value of L for kernel estimation, where the input of the prediction is the given blurred image B.

In the prediction step, we compute gradient maps of L along the x and y directions which predict salient edges in L with noise suppression in smooth regions. Except at the beginning of the iteration, the input of the prediction step is the estimate of L obtained in the deconvolution step of the previous iteration. In the kernel estimation step, we estimate K using the predicted gradient maps of L and the gradient maps of B. In the deconvolution step, we obtain an estimate of L using K and B, which will be processed by the prediction step of the next iteration.

To make the estimations of K and L more effective and efficient, our method employs a **coarse-to-fine scheme(Fig. 5)**. At the coarsest level, we use the down-sampled version of B to initialize the process with the prediction step. After the final estimate of L has been obtained at a coarse level, it is up-sampled by bilinear interpolation and then used for the input of the first prediction step at the next finer level. In our experiments, we performed seven iterations of the three steps at each scale. As detailed in Sec. 4, this coarse-to-fine scheme enables handling of large blurs for which prediction with image filtering may not suffice to capture sharp edges.

In the **coarse-to-fine** iterative process for updating K and L, we use the gray-scale versions of B and L. After the final K has been obtained at the finest level, i.e., with the input image size, we perform the final deconvolution with K on each color channel of B to obtain the deblurring result. Although any non-blind deconvolution technique can be used for this step, we use the technique proposed in [Shan et al. 2008], which efficiently obtains high quality deconvolution results. Fig. 5 shows an example of our deblurring process with intermediate estimates of K and L.

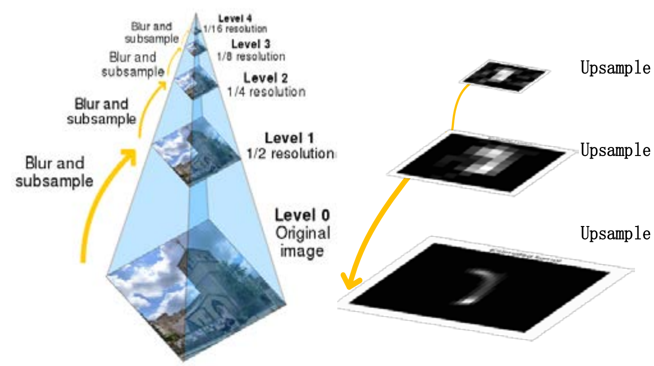


Fig. 5. Estimated kernels at different scales for the deblurring example.

## *E.* Large blurs

Our prediction method may fail to correctly predict sharp edges for large blurs. However, our **coarse-to-fine** scheme enables us to avoid direct prediction of edges from a largely blurred image. We first predict sharp edges in a low resolution image,where the extents of blurs have been narrowed and most edges can be predicted without severe localization errors. At a higher resolution, we start prediction of sharp edges with an upsampled version of the deconvolved image obtained at a coarser resolution, which contains reduced amounts of blurs. In the iterative process at a specific scale, sharp edge prediction is applied to the deconvolution result obtained by an updated kernel in the previous iteration, progressively improving the prediction accuracy. With the multi-scale iterative process, we can estimate kernels for large blurs using prediction with small size bilateral and shock filters. This multi-scale iterative process is also the main difference from [Joshi et al. 2008] and [Money and Kang 2008], which enables our method to estimate complex large motion blurs that cannot be handled by the previous methods. **Fig. 5 shows estimated kernels at different scales in our multi-scale process.**

## IV. Results

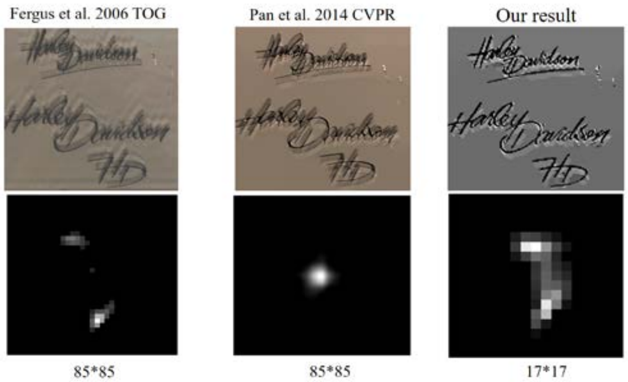


Fig. 6. Compare Fergus's , Pan's and our results

Fig. 6 shows our deblurring results of real photographs. The text photographs contain complex structures and different camera motions. In the deblurring results, sharp edges also have been significantly enhanced, revealing the object shapes and structures more clearly.

## V. Conclusion

In this paper, we propose a simple yet effective prior for text image deblurring. We discuss how the proposed prior facilitates preserving salient edges in image deblurring, and extend it to deal with natural images. With this prior, we present an effective optimization algorithm based on the half-quadratic splitting approach, which ensures that each sub-problem has a closed-form solution. Experimental results show that the proposed algorithms perform favorably against the state-of-the-art methods for deblurring text images without additional pre-processing steps. In addition, we develop a latent image restoration method which helps reduce artifacts effectively. The proposed algorithm is also extended to deblur natural images and low-illumination scenes, as well as non-uniform cases.

**Sensibility:**

We observed that our deblurring results are relatively sensitive to parameters, and improving the robustness to the parameters is our future work. Nevertheless, we could obtain desirable results usually within a few trials, where even several trials need much less time than a single run of previous methods.

**Limitation:**

The proposed methods are likely to fail when a blurred image contains a large amount of noise (e.g., Gaussian and non-Gaussian noise) as the data term used in the proposed model is based on L2 norm which is less robust to noise. In addition, the proposed intensity prior counts the number of pixels with nonzero-intensity values. As all the pixels are treated independently, the proposed algorithm is sensitive to large image noise and exacerbates its effect in the Pan's results (See Fig. 6.)

**Future work:**

Pan's method shares common limitations with other uniform motion deblurring methods. Due to the limitation of the blur model based on convolution, saturated pixels from strong lights, **severe noise**, and a spatially varying blur would not be properly handled. Extending our deblurring method to resolve this limitation will be interesting future work.

Our future work can focus on simultaneously denoising and deblurring for both text and natural images with severe noise.

## References


[1] Richardson, William Hadley. "Bayesian-based iterative method of image restoration." *JOSA* 62.1 (1972): 55-59.

[2] Lucy, L. B. "An iterative technique for the rectification of observed distributions."*The astronomical journal* 79 (1974): 745.

[3] Wiener, Norbert. *Extrapolation, interpolation, and smoothing of stationary time series: with engineering applications*. Technology Press of the Massachusetts Institute of Technology, 1950.

[4] J. Bardsley, S. Jefferies, J. Nagy, and R. Plemmons. A computational method for the restoration of images with an unknown, spatially-varying blur. Optics Express, 14(5):1767–1782, 2006. 2

[5] M. Ben-Ezra and S. Nayar. Motion deblurring using hybrid imaging. In CVPR, pages 657–664, 2003. 2

[6] A. Chakrabarti, T. Zickler, and W. T. Freeman. Analyzing spatially-varying blur. In CVPR, pages 2512–2519, 2010. 2

[7] Q. Chen, D. Li, and C.-K. Tang. Knn matting. In CVPR, pages 869–876, 2012. 4

[8] S. Cho and S. Lee. Fast motion deblurring. In SIGGRAPH Asia, 2009. 2, 4, 6, 8

[9] S. Cho., Y. Matsushita, and S. Lee. Removing non-uniform motion blur from images. In ICCV, 2007. 1, 2

[10] T. S. Cho, N. Joshi, C. L. Zitnick, S. B. Kang, R. Szeliski, and W. T. Freeman. A content-aware image prior. In CVPR, pages 169–176, 2010. 2

[11] Banerjee, J., Namboodiri, A., Jawahar, C.: Contextual restoration of severely degraded document images. In: Proc. CVPR 2009, pp. 517–524 (2009)

[12] Li, T.H., Lii, K.S.: A joint estimation approach for two-tone image deblurring by blind deconvolution. IEEE Trans. Image Processing 11, 847–858 (2002)

[13] Donaldson, K., Myers, G.: Bayesian super-resolution of text in video with a text-specific bimodal prior. In: Proc. CVPR 2005, vol. 1, pp. 1188–1195 (2005)

[14] Lam, E.: Blind bi-level image restoration with iterated quadratic programming. IEEE Trans. Circuits and Systems II: Express Briefs 54, 52–56 (2007)

[15] Cho H , Wang J , Lee S . Text Image Deblurring Using Text-Specific Properties[C]// European Conference on Computer Vision. Springer, Berlin, Heidelberg, 2012.

[16] Pan J , Hu Z , Su Z , et al. $L_0$ -Regularized Intensity and Gradient Prior for Deblurring Text Images and Beyond[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39(2):342-355.