SINGLE IMAGE DEBLURRING WITH ADAPTIVE DICTIONARY LEARNING

Zhe Hu   Jia-Bin Huang   Ming-Hsuan Yang

Electrical Engineering and Computer Science
University of California at Merced
Merced, CA 95344, USA

ABSTRACT

We propose a motion deblurring algorithm that exploits sparsity constraints of image patches using one single frame. In our formulation, each image patch is encoded with sparse coefficients using an over-complete dictionary. The sparsity constraints facilitate recovering the latent image without solving an ill-posed deconvolution problem. In addition, the dictionary is learned and updated directly from one single frame without using additional images. The proposed method iteratively utilizes sparsity constraints to recover latent image, estimates the deblur kernel, and updates the dictionary directly from one single image. The final deblurred image is then recovered once the deblur kernel is estimated using our method. Experiments show that the proposed algorithm achieves favorable results against the state-of-the-art methods.

Index Terms— Image deblurring, sparse representation, blind deconvolution.

1. INTRODUCTION

Image deblurring is one of the most fundamental problems in image restoration that has been studied extensively in the literature [3, 2]. Image blur is usually caused by relative motion between the camera and the scene during the exposure time, e.g., camera shake. When there is a lack of ambient light, slower shutter speed is necessary to increase exposure time. Consequently, camera shake is likely to happen and degradate the image quality significantly. The image blur caused by camera shake can be modeled by a latent image convolving with a spatial-invariant kernel K:

$$B = K \otimes I + n, \qquad (1)$$

where B, I and n represent the input blurry image, deblurred (or latent) image as well as noise respectively. Within this formulation, $\otimes$ denotes the convolution operator and the deblurring problem is thus posed as a deconvolution problem.

If the kernel K is assumed to be known or well estimated, it becomes a non-blind deconvolution problem. The Weiner filter [13] and Richardson-Lucy algorithm [11, 10] are two widely used non-blind deconvolution methods due to their simplicity and efficiency. Nevertheless, these two methods often introduce noticeable ringing effects. In [9], Lou et al. propose a method that operates directly on blurred images, without deconvolution, using a set of learned basis and a *known* kernel. Their intuition is that, given a set of learned image basis, the sparse coefficients of a latent image I remain the same no matter whether I has been blurred or not. Instead of tackling image deblurring with deconvolution, one can convolve the image basis with K from which sparse coding of the blurred image I can be computed. However, in most of the real-world cases, we do not know the exact kernel for a blurred image, and the deblurring problem with unknown kernel is posited as a blind-deconvolution problem.

Blind deconvolution is a well-known ill-posed problem, and deblurring methods within this formulation aim at recovering images with missing high frequency details. To address this problem, some methods first estimate the kernel K and then employ non-blind deconvolution algorithm to obtain the latent image I [7, 14], while others estimate kernel K and latent image I simultaneously [12, 4]. With known or well-estimated kernel, there exist numerous algorithms that can recover the blurred image well (e.g., [11, 10]). If the kernel can be estimated, we can employ these non-blind deconvolution methods to obtain high-quality latent image. Prior gradient distribution of nature images and sophisticated machine learning algorithm are utilized for kernel estimation in [7]. However, rich and specific information contained in a blurred image is not exploited for estimating deblur kernel. In [14] a method that utilizes additional information from a pair of blurred and noisy images is proposed to estimate kernel. Most recently, an algorithm that models the deblurred blur kernel and image with wavelets and sparsity constraints has demonstrated promising results [4]. Nevertheless, this method does not exploit rich information contained in the blurred image as the generic bases (i.e., pre-defined basis such as curvelets and framelets) are used.

In this paper, we propose a deblurring method that exploits sparse representation using a dictionary learned directly from the blurred image. Our method utilizes sparsity constraints to iteratively estimate the deblur kernel and then applies a standard non-blind deconvolution algorithm to recover the deblurred image. While other methods utilizes prior statistics learned from a set of additional images for deblurring [7, 8, 14, 12], the proposed algorithm requires only one image. In contrast to [4], our method benefits from dictionary

learning from the blurred image itself. Compared with state-of-the-art methods, our experiments show that the proposed deblurring algorithm achieves favorable results.

2. PROBLEM FORMULATION

It is well-known that natural images can be modeled with sparse representation over an over-complete dictionary. A signal (an image patch in this paper) I_p can be approximated by a sparse linear combination of atom signals in an over-complete dictionary $D \in \mathcal{R}^{n \times m}$, $n, m \in \mathcal{R}$ and $n < m$, or a set of overcomplete bases $\{d_j\}_{j=1}^m$ and a coefficient vector $\alpha \in \mathcal{R}^m$, that satisfy:

$$\min \|\alpha\|_0, \quad \text{s.t.} \quad I_p = D\alpha, \tag{2}$$

$$\text{or} \quad \min \|\alpha\|_0, \quad \text{s.t.} \quad \|I_p - D\alpha\|_2 \leq \epsilon, \tag{3}$$

where I_p is an image patch that has the same dimension with the dictionary basis and $\| \cdot \|_k$ is the l_k norm. In other words, any signal can be described using a small number of dictionary bases under the corresponding bases pool.

We pose image deblurring as a joint optimization problem of the blur kernel K and sparse coefficients α with regularization:

$$\min_{K,\alpha} \|B - K \otimes DA\|_2^2 + \lambda \sum \Theta(\alpha), \tag{4}$$

where each column of B is a raster scan vector of a block from the blurred image, and each column of A is the sparse coefficient α of a block of the latent image (from (2) or (3)). In addition, $\Theta(\alpha)$ is the l_1 regularization term upon the sparse representation coefficient α as it has been proved to be equivalent to l_0 norm under certain conditions [6]. In (4), there are three unknown variables: the kernel K, the coefficient vector α, and the underlying dictionary D. Therefore, a direct approach to solve this optimization problem is likely to suffer from slow convergence and local minimums. As the main goal is to estimate the deblur kernel, we propose an iterative method to alternately estimate the unknown variables, one at a time, which divides the optimization problem into several simple ones in each iteration. More importantly, the dictionary D is learned from the input image during this optimization process. The proposed algorithm iteratively optimizes one of K, D, α by fixing the other two, and finally obtains the deblurring kernel. With the estimated kernel, we can apply any standard deconvolution algorithm to recover the latent image.

2.1. Estimating sparse coefficients

At the beginning of each iteration, we fix K and D to estimate the coefficient α of each image patch, the column of A

$$\alpha^{(n+1)} = \arg\min \|\alpha\|_1, \quad \text{s.t.} \quad b = (K^{(n)} \otimes D^{(n)})\alpha,$$

$$\text{or} \quad \alpha^{(n+1)} = \arg\min \|\alpha\|_1, \quad \text{s.t.} \quad b = \hat{D}^{(n)}\alpha,$$
$$\tag{5}$$

where $\hat{D}^{(n)}$ is the blurred dictionary at iteration n. Here, we use b to represent a block of blur image. At the first iteration,

we need to initialize K_0, which will be discussed in Sec. 3. As K and D are fixed at iteration n, there is only one unknown variable and the optimization problem can be solved efficiently using the basis pursuit algorithm [5].

2.2. Updating dictionary

In this step, as kernel K and coefficients α are fixed, we update the dictionary atoms of the latent image based on the test image itself. However, the extracted information from the blurred image B cannot be directly applied to update the underlying dictionary of the clear image. Before the update of dictionary D, we reconstruct an image $\hat{I}$ from deconvolving the blurred image with the estimated kernel $K^{(n)}$ using a non-blind deconvolution algorithm, such as the Richardson-Lucy or Wiener deconvolution algorithms, and then use the blocks $\hat{I}_p$ of the image, which shares the same dimension as the dictionary basis, to update the dictionary D of latent image. Similar to the K-SVD algorithm [1], we decompose the multiplication $D\alpha$ to the sum of rank-1 matrices and each time update a basis d_i of dictionary D:

$$\begin{aligned} d_i^{(n+1)} &= \min \|\hat{I}_p - D^{(n)}\alpha^{(n+1)}\|_2^2 \\ &= \min_{d_i} \|\hat{I}_p - (d_i\alpha_i^{n+1} + \sum_{j \neq i} d_j^{(n)}\alpha_j^{(n+1)})\|_2^2 \\ &= \min_{d_i} \|E_i^{(n+1)} - d_i\alpha_i^{(n+1)}\|_2^2, \end{aligned} \tag{6}$$

where $E_i^{(n+1)} = \hat{I}_p - \sum_{j \neq i} d_j^{(n)}\alpha_j^{(n+1)}$ is the residual matrix.

2.3. Recovering latent image and estimating blur kernel

Similar to the ideas proposed in [9], we assume the sparse coefficients α of the latent image patch I_p with basis $\{d_i\}_{i=1}^m$ are the same as the one of blurred image B relative to the blurred basis $\{b_i\} = \{K \otimes d_i\}_{i=1}^m$. Thus, we can reconstruct latent image $I^{(n+1)}$ as the weighted sum of overlapping image blocks

$$I_p^{(n+1)} = D^{(n+1)}\alpha^{(n+1)}. \tag{7}$$

Next, we can write the convolution $K^{(n+1)} \otimes I^{(n+1)}$ as a matrix multiplication $A \cdot x$ (here x corresponds to $K^{(n+1)}$ and image blocks I_p as the columns of A) and formulate the kernel estimation problem as

$$\min \|B - Ax\|_2^2. \tag{8}$$

Moreover, we introduce additional regularization constraints in order to get a stable solution from Tikhonov regularization,

$$\min \|B - Ax\|_2^2 + \|\Gamma x\|_2^2, \tag{9}$$

where Γ is the Tikhonov matrix. With Tikhonov regularization, the numerical solution, denoted by $\hat{x}$, is given by

$$\hat{x} = (A^\top A + \Gamma^\top \Gamma)^{-1} A^\top B. \tag{10}$$

Alternatively, we can use objective functions with penalty of l_1 norm [12] or framelet system [4], such as

$$\min \|B - Ax\|_2^2 + \|x\|_1, \tag{11}$$

to estimate K. As mentioned above, the final deblurred image can be recovered once the deblur kernel is estimated.

2.4. Algorithm analysis

Our algorithm focuses on estimating the blur kernel and then employs a standard non-blind deconvolution algorithm to recover the latent image. The success of our algorithm hinges on the effectiveness and accuracy of kernel estimation. That is, if the accuracy of the estimated kernel is improved at each iteration, the proposed algorithm will find a reasonably good solution. This condition can be achieved under certain situations. For minimization problem (5), we have $\hat{D} = K_e \otimes D_I$, where K_e and D_I represent the estimated kernel and dictionary for the latent image I, and $K_r = K_e + E$, where K_r is the real kernel and E is the kernel error. If this minimization problem has been solved precisely, $B = (K_r - E) \otimes D_I \alpha$. Meanwhile, $B = K_r \otimes I$, then

$$
\begin{aligned}
K_r \otimes I &= (K_r - E) \otimes D_I \alpha \\
\implies K_r \otimes (I - D_I \alpha) &= -E \otimes D_I \alpha,
\end{aligned} \tag{12}
$$

Based on the assumption of coefficients α stated above, if the dictionary D_I for latent image is correctly estimated ($D_I \alpha \approx I$), which means that the error $\|I - D_I \alpha\|_2$ will be decreased. Although both sides of the above equation have convolution operations, from the definition of convolution, it is reasonable to expect that the variation trend of $\|E\|_2$ is being decreased. That is, the estimate of blur kernel is improved over each iteration and thus the latent image can be recovered.

3. EXPERIMENTAL RESULTS

We implement our algorithm and carry out experiments in MATLAB. The initial kernel K_0 is set to be the Gaussian kernel with $\sigma = 1$, and λ in (5) as well as Γ in (10) are set as 1 and identity matrix I. We use gray-scale images for our experiments and crop a small portion (e.g. 200×200 pixels) of the tested image to estimate kernel using the algorithm listed in Sec. 2. We use Richardson-Lucy deconvolution algorithm to reconstruct $\hat{I}$ in Sec. 2.2. Each iteration of our algorithm takes approximately 5 minutes on a windows PC of 2.67 GHz CPU and 4 GB RAM.

We compare our method with two state-of-art algorithms [7, 12] which utilize prior information on image edges. As shown in Fig. 1, although the deblurred images from [7, 12] have shaper edges, they also contain noise in smooth regions such as walls and floors. In contrast, our algorithm is able to recover the deblurred images well in smooth and textured regions (e.g., fluff of the koala, stipes of tablecloth and kerchief) with fewer ringing effects. One possible explanation is that the assumed prior distribution of edges used in [7, 12] does not hold well for these images whereas our method exploits the information from the blurred image.

We also carry out experiments with images blurred by randomly generated kernel. The existing deblurring algorithms are usually developed to deal with motion blur problems in which the kernels are oriented and simple. However, the camera shakes are complex and cannot be modeled well with simple blur kernels. We evaluate our algorithms against others using randomly generated kernels. Fig. 2 show that our algorithm is able to recover the latent image with more details and better contrast than [7] (which does not capture image details) and [12] (which suffers more ringing effects with fewer details). We also provide root-mean-square-error (RMSE) to demonstrate the quantitative comparison among recovered images by different algorithms. These results can be better viewed with zoom-in on a LCD display, and their high-resolution images can be found at eng.ucmerced.edu/people/zhu/deblur.html.

4. CONCLUSION

In this paper, we propose an effective deblurring algorithm with adaptive dictionary learning using one single image. By decomposing the blind deconvolution problem into three simple ones and learning sparse dictionary from the image, our method is able to estimate blur kernels and thereby deblurred images. Experimental results show that our algorithm achieves favorable performance against two state-of-the-art algorithms.

5. REFERENCES

[1] M. Aharon, M. Elad, and A. Bruckstein. K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Transactions on Signal Processing*, 54(11):4311–4322, 2006.

[2] M. Ben-Ezra and S. Nayar. Motion-based motion deblurring. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(6):689–698, 2004.

[3] J. Biemond, R. Lagendijk, and R. Mersereau. Iterative methods for image deblurring. *Proceedings of the IEEE*, 78(5):856–883, 1990.

[4] J. Cai, H. Ji, C. Liu, and Z. Shen. Blind motion deblurring from a single image using sparse approximation. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 104–111, 2009.

[5] S. S. Chen, D. L. Donoho, and M. A. Saunders. Atomic decomposition by basis pursuit. *SIAM Review*, 43(1):129–159, 2001.

[6] D. L. Donoho. For most large underdetermined systems of linear equations the minimal l_1-norm solution is also the sparsest solutions. *Communications on Pure and Applied Mathematics*, pages 797–829, 2004.

[7] R. Fergus, B. Singh, A. Hertzmann, S. T. Rowels, and W. T. Freeman. Removing camera shake from a single photograph. In *SIGGRAPH*, 2006.

[8] A. Levin. Blind motion deblurring using image statistics. In *Advances in Neural Information Processing Systems*, pages 841–848, 2006.

[9] Y. Lou, A. Bertozzi, and S. Soatto. Direct sparse deblurring. Technical report, UCLA CAM report 09-15, 2009.

[10] L. Lucy. An iterative technique for the rectification of observed distributions. *Astronomical Journal*, 79(6):745–754, 1974.

[11] W. Richardson. Bayesian-based iterative method of image restoration. *Journal of the Optical Society of America*, 62(1):55–59, 1972.

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

[13] N. Wiener. *Extrapolation, Interpolation and Smoothing of Stationary Time Series*. MIT Press, 1964.

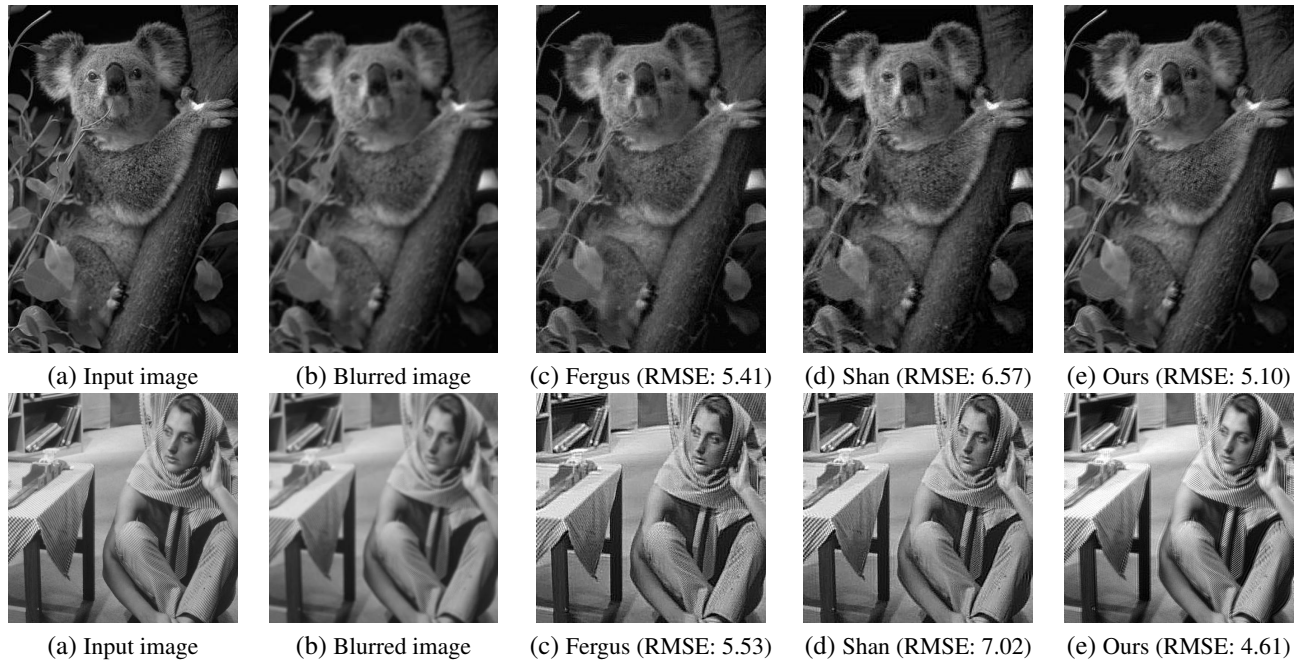[14] L. Yuan, J. Sun, L. Quan, and H. Shum. Image deblurring with blurred/noisy image pairs. In *SIGGRAPH*, 2007.

(a) Input image   (b) Blurred image   (c) Fergus (RMSE: 5.41)   (d) Shan (RMSE: 6.57)   (e) Ours (RMSE: 5.10)

(a) Input image   (b) Blurred image   (c) Fergus (RMSE: 5.53)   (d) Shan (RMSE: 7.02)   (e) Ours (RMSE: 4.61)

Fig. 1. Experiments with motion blur kernel. (a): input image; (b): blurred image; (c): recovered images using approach by Fergus et al. [7]; (d): recovered images using approach by Shan et al. [12]; (e): recovered images by our algorithm.

(a) Input image   (b) Blurred image   (c) Fergus (RMSE: 7.87)   (d) Shan (RMSE: 7.46)   (e) Ours (RMSE: 6.73)

(a) Input image   (b) Blurred image   (c) Fergus (RMSE: 6.58)   (d) Shan (RMSE: 7.21)   (e) Ours (RMSE: 6.94)

Fig. 2. Experiments with randomly generated kernel. (a): input image; (b): blurred image; (c): recovered images using approach by Fergus et al. [7]; (d): recovered images using approach by Shan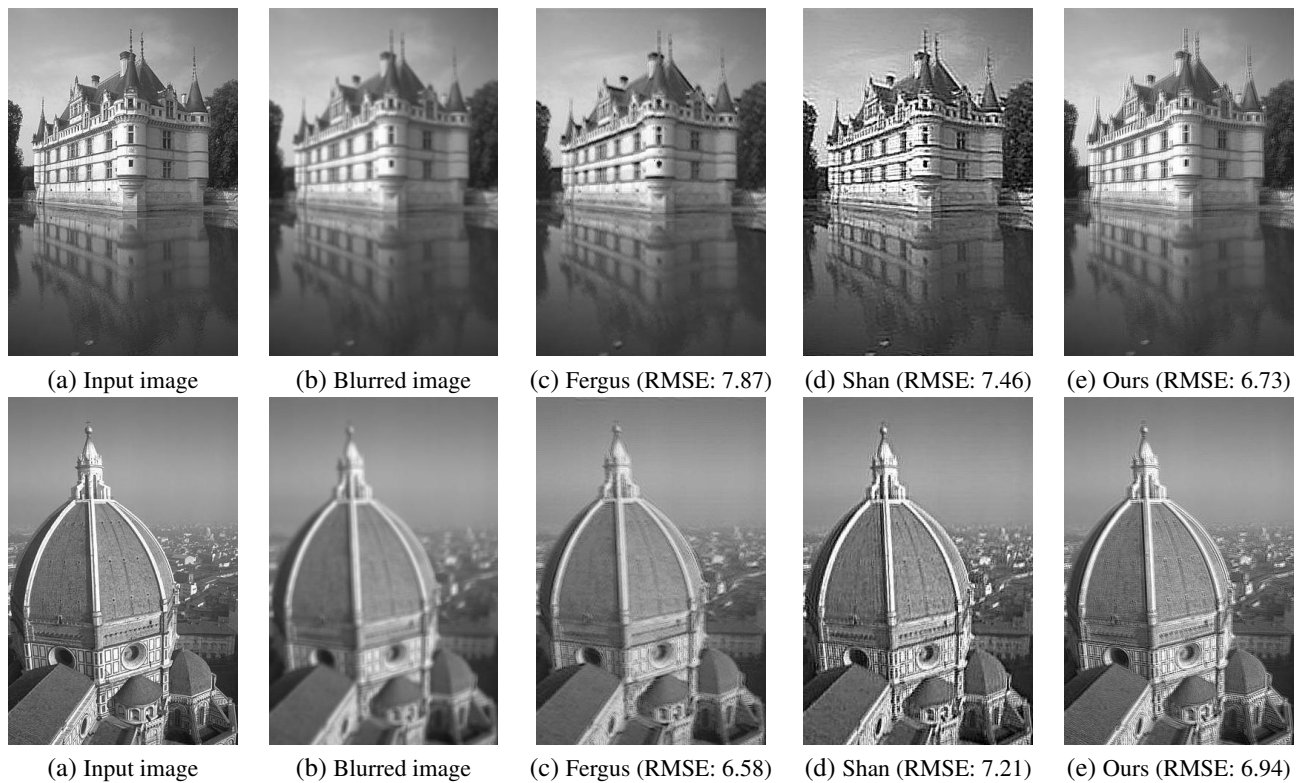 et al. [12]; (e): recovered images by our algorithm.