REMOVING OUT-OF-FOCUS BLUR FROM SIMILAR IMAGE PAIRS

Jiangyong Duan, Gaofeng Meng, Shiming Xiang, Chunhong Pan

NLPR, Institution of Automation, Chinese Academy of Sciences, China
{jyduan, gfmeng, smxiang, chpan}@nlpr.ia.ac.cn

ABSTRACT

This paper presents a new deblurring method to remove the out-of-focus blur from similar image pairs. The method is motivated by an observation that a blurred structure appearing in one image can often have its corresponding clear one in the similar clear images. Our method first extracts the patch pairs from input images by SIFT matching. Then the constraints on the patch pairs are used to estimate the blur kernel via the RANSAC algorithm. Finally, the non-blind deconvolution is adopted to restore the blurred image. The main advantage is that we can improve the deblurring results with the help of additional similar clear images in many practical applications. Our method is validated on synthetic and real images by comparing with state-of-the-art methods.

***Index Terms*—** Image deblurring, out-of-focus blur, SIFT matching, region pairs

1. INTRODUCTION

When taking photos, one general problem is the blurs in the captured images. Caused by cameras shaking or objects out-of-focus, these blurs significantly degenerate the visual qualities of the images. In the last decades, many image deblurring methods have been proposed to restore the clear images. In these methods, the image deblurring is generally modeled as a problem of blind deconvolution, in which the blurred image is generated by a convolution of a clear image with a blur kernel. Due to the unknown clear image and kernel, image deblurring is a challenging ill-posed problem.

To make the problem well defined, one kind of methods adopt the statistical priors learned from the natural images. Levin [1] assumes that the gradients of natural images follow a Laplacian distribution. This prior is later used to constrain the estimation of a motion blur kernel. Fergus and Singh [2] combine the prior of the image gradients with the prior of noise to maximize the posterior probability of the kernel. Shan et al. [3] go a further step by using the priors of image, kernel and noise to maximize the joint probability of the latent image and the kernel. Other priors, e.g., gradient Hyper-Laplacian prior [4], normalized sparse prior [5] and variant gradient priors [6], are also introduced in recent methods. Although great improvements have been achieved by these methods, the statistical priors alone are not adequate for the deblurring.

To improve the generic model derived from the statistical priors of a single image, the multi-image based methods are developed. Rav and Peleg [7] use the constraints between two different motion blurred images to estimate the motion kernel. Yuan et al. [8] take a noisy image as the latent image of the blurred image to compute the blur kernel. Zhuo et al. [9] use a no-flashed image to deblur a flashed motion blurred image. In addition, to avoid estimating the blur kernel, Xiang et al. employ blurred-clear image pairs as samples to train matrix mapping functions in terms of supervised learning [10]. Although satisfied results are produced, their experimental settings make them difficult to be applied in practical situations. For example, two aligned images for the same scene are required in [8, 9].

Other methods are also developed with the help of additional hardware. Levin et al. [11] use a coded aperture to estimate the depth of the scene and then recover the all in-focus image by deconvolution. Joshi et al. [12] use an inertial measurement sensors to record the acceleration and angular velocity during the motion blur and then estimate the blur kernel. Although high-quality results are obtained, their methods can not be applied for traditional cameras.

In many applications, we can easily obtain several clear images with similar scenes to the blurred images. The objects defocused in one image will be clear in the other images when they are focused. One example is illustrated in Fig. 1. The flowers and the leaves blurred in one image are clear in another image. In such cases, the clear images can provide useful information to remove the blur in the blurred images. Specifically, the patches in the blurred images can have their corresponding clear ones in the other images.

Fig. 1. Two images taken at different focuses. While some of the regions in one image are blurred, their corresponding regions in another image are clear. The blurred and clear regions are bounded separately by red and green rectangles.

In this paper, we propose a method to remove the out-of-focus blur from similar clear images. It arises from the observation that a patch in a blurred image can often have a corresponding clear one in their similar images. Our method first extracts the patch pairs by using SIFT matching. Then, we estimate the blur kernel by the constraint between the patch pairs via the RANSAC algorithm. Finally, the clear image is restored by image deconvolution.

The main contribution of our method is that we adopt new constraints extracted from similar clear images for the deblurring. Compared with methods [1-6], our constraints are more specific than the statistical priors. While the statistical priors are generic for all the natural images, our constraints are specifical to the blurred image. In fact, our constraints can be a complementary to the statistical priors. In compared with methods with multiple images [7-9], our method relaxes the experimental requirements in their methods, i.e.,

This work was supported in part by the National Natural Science Foundation of China under Grants 61005036, 61175025, 61005013 and 61272331.

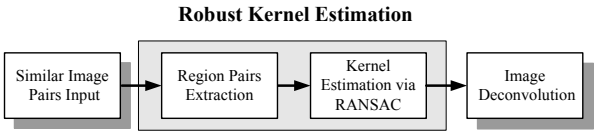

Fig. 2. The framework of our algorithm

the aligned images for the same scene are replaced by the unaligned images for similar scene. Finally, our methods can be applied for the images captured by either the traditional cameras or the computational camera with the coded aperture [11].

The reminder of the paper is organized as follows: Section 2 describes our method. The experiments are conducted and analyzed in Section 3. Finally, we conclude our paper in Section 4.

2. OUR ALGORITHM

The goal of our method is to remove the out-of-focus blur. This blur, namely the Bokeh effect, happens when parts of the scene lie outside the depth of field. In our method, we adopt the common assumption used in existing methods [1, 2] that the blur is shift-invariant. Under this assumption, the out-of-focus image can be modeled as a convolution of a latent clear image with a blur kernel:

$$B = I \otimes K, \tag{1}$$

where B is the blurred image, I is the latent clear image and K is the blur kernel, $\otimes$ is the convolution operator. Although the assumption is not true for the out-of-focus image, it is valid for some of the regions which have the approximately same depths. However, it is still a challenging problem to segment the scene with the same depth [13]. In this paper, we manually select these regions to be deblurred.

In (1), the problem is ill-posed because both the latent image I and the blur kernel K are unknown. Fortunately, we can easily obtain the similar clear images along with the out-of-focus image as shown in Fig. 1. In these similar clear images, they often contains the clear regions corresponding to the blurred regions in the out-of-focus image. This inspires us to use the constraints between the region pairs to remove the blur, where the region pairs can be extracted from the out-of-focus image and its similar clear images.

Our algorithm first extracts region pairs. This is implemented by extracting and matching SIFT points in input images. Then the image pairs are aligned by affine transformation and the region pairs are extracted around each the matching point pair in each input image. With the region pairs, the constraints are obtained and the blur kernel is estimated by enforcing these constraints via the RANSAC algorithm. Finally, the blurred image is restored by deconvolution. The framework of our algorithm is illustrated in Fig. 2 and an example is shown in Fig. 3.

2.1. Region Pairs Constraints

Given R^B, a blurred region from the blurred image B, and R^{I_S}, its corresponding clear region in the clear image I_S, we have the following constraint between the region pair

$$R^B = R^{I_S} \otimes K. \tag{2}$$

To extract the region pairs, we propose a method based on SIFT matching [14]. The reason is explained below. Generally, the out-of-focus image can be approximately considered as the convolution of an latent clear image with a Gaussian kernel [15]. Consequently,

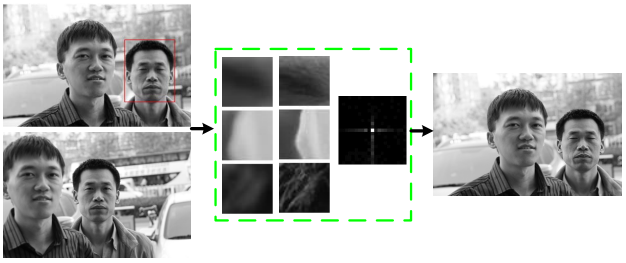

Fig. 3. An example of our deblurring framework: (a) input image pairs. The region in red box is to be deblurred, (b) the extracted region pairs and the estimated kernel, (c) the restored image.

the out-of-focus image resembles the scaled image of the latent clear image. Therefore, the SIFT matching can extract the corresponding scale-invariant points in the input images.

As described above, the point pairs are first automatically extracted and matched by SIFT matching. Then, affine transform is used to align the similar clear image to the blurred image. After alignment, the $N * N$ region pairs $\{R_i^B, R_i^{I_S}\}$ are obtained by extracting the regions around the point pairs. For the extracted region pairs, the regions in the similar clear image will correspond to the regions in the blurred image. Fig. 4 illustrates some of the extracted region pairs from the images in Fig. 3. As shown in Fig. 4, the extracted region pairs are indeed the corresponding regions in the out-of-focus image and similar clear image.

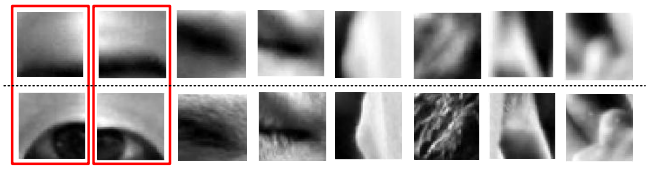

Fig. 4. Extracted region pairs: the first row shows the regions in the out-of-focus image, the second row shows the corresponding clear regions in the similar clear image. The red rectangles show some outliers in the extracted region pairs.

After the patch pairs extraction, one simple deblurring method is to replace the blurred image by its aligned clear images. Unfortunately, this method often fails. For example, the eyes of the man in the restored image in Fig. 3 will be open by direct substitution. However, the eyes of the man in the blurred image should be closed. Instead, we adopt an alternative method by estimating the blur kernel using constraint between the region pairs and then restore the image by deconvolution.

2.2. Robust Kernel Estimation

In (2), there is a linear constraint for the blur kernel K on each region pair which is extracted from each of the multiple input similar clear image. With some sets of region pairs obtained, K are usually over-constrained and can be easily estimated. Actually, we can generally extract more than twenty region pairs in our experiments, which are adequate for estimating the blur kernel.

Our objective function with respective to K is defined as

$$L(K) = \sum_{i=1}^{M} ||R_i^B - R_i^{I_S} \otimes K||_2^2, \tag{3}$$

where $\{R_i^B, R_i^{I_S}\}$ are the extracted region pairs, M is the number of the extracted pairs. By minimizing (3), K is uniquely estimated and has a closed-form solution in the frequency domain:

$$K = F^{-1}\left\{\frac{\sum_{i=1}^{M} F^*(R_i^{I_S})F(R_i^B)}{\sum_{i=1}^{M} F^*(R_i^{I_S})F(R_i^{I_S})}\right\}, \quad (4)$$

where F and F^{-1} are the Fourier transform and the inverse Fourier transform respectively, and the superscript $*$ represents the conjugate of the complex.

However, the result of (4) will be deviate from the true kernel when some region pairs do not satisfy (2). Fig. 4 shows two outlier pairs. In these two region pairs, the eyes in the blurred image are closed, while the clear eyes in the similar clear image are open. Therefore, the estimated kernel by these pairs will be inaccurate . To this end, robust method RANSAC [16] is used to remove these outliers. It proceeds as follows:

1. Randomly select a region pair $\{R_i^B, R_i^{I_S}\}$ to compute the initial kernel K_i [17]:

$$K_i = F^{-1}\left\{\frac{F^*(R_i^{I_S})F(R_i^B)}{F^*(R_i^{I_S})F(R_i^{I_S})}\right\}; \quad (5)$$

2. Use K_i estimate the inlier region pairs. We select the region pairs with the small errors:

$$E_j^i = ||R_j^B - R_j^{I_S} \otimes K_i||_2^2, \quad j = 1, \ldots, M; \quad (6)$$

3. Re-estimate the kernel $\bar{K}_i$ using all the inlier pairs, and compute the error of $\bar{K}_i$ by summing all the errors of the inlier region pairs;
4. Go through all the region pairs and choose the kernel with the minimal error as the final kernel.

2.3. Deconvolution

After the kernel has been estimated, we use Richardson-Lucy deconvolution method [18] to restore the image. It iteratively computes the latent image I until convergence:

$$I^{t+1} = I^t \times \tilde{K} \otimes \frac{B}{K \otimes I^t}, \quad (7)$$

where $\tilde{K}$ is the transpose of K that flips the shape of K upside-down and left-to-right, and $\times$ is the pointwise product operator. While it is simple, we can produce satisfied results because of the accurately estimated blur kernel.

3. EXPERIMENTS

We validate our algorithm on synthetic and real images. Three state-of-the-art methods, i.e., Shan et al. [3][1], Li and Jia [17][2] and Krishnan et al. [5][3], are compared with our method. The codes are provided by the authors and the parameters in the codes are tuned to produce the best results.

In our method, the extracted patch size is larger than the kernel size to ensure (5) can be computed. In our experiments, we set it to 41×41. The number of the inliers in RANSAC is set to be the half of the number of the region pairs. The size of the blur kernel is determined by the level of the blur and 17×17 is used by default.

[1] http://www.cse.cuhk.edu.hk/∼leojia/projects/motion_deblurring/

[2] http://www.cse.cuhk.edu.hk/∼leojia/deblurring.htm

[3] http://cs.nyu.edu/∼dilip/research/blind-deconvolution/

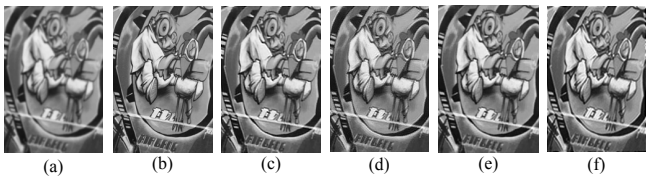

Fig. 5. Comparisons on synthetic images. (a) Blurred image, (b) ground truth (also used as the similar clear image in our method), (c)-(f) results of [17], [3], [5] and ours respectively.

For synthetic images, we evaluate the Mean Square Error (MSE) between the restored image and the ground truth. The results are illustrated in Fig. 5. Our method restores the clear image, in which the blurred edges are recovered and the characters in the blurred image can be recognized. The MSE of [17], [3], [5] and ours are 16.42, 18.57, 16.57 and 17.60 respectively. Our MSE is comparable to the MSE of the other methods. The visual and quantity comparisons indicate that our constraints on region pairs indeed help the deblurring and can play the role of statistical prior in the other methods.

The comparisons on real images are illustrated in Fig. 6. While our method produces the comparable results on the overall blurred region, we obtain the better results in recovering more details and produce fewer artifacts. This is shown in the restored eye in Fig. 6. In [3], obvious artifacts of blocks appear around the eye. In [17] and [5], the details are lost and the restored image are smoothed in the eyelid. However, our method produces the satisfied results with the more details and fewer artifacts. While [17] and [5] seams clear in that the edges of the restored images are enhanced, our result is more natural because both the texture and the details are restored.

We show another comparative results on real image in Fig. 7. As shown, the results of [3] and [17] are much more blurred than the results of [5] and ours. Whereas, the method in [5] tends to smooth the image in the textural regions and thereby loses many details around the eye in Fig. 7 (e). Contrast to [5], our result contains more details and the restored eye is more human-like because the skin of the eye has fine texture rather than just a smoothed surface.

The reason that better results are produced is that we estimate the accurate blur kernel with additional similar clear images. For the other three methods, statistical priors are adopted to estimate the blur kernel. These priors are generic for all the natural images. Contrarily, the constraints in our method is specific to the blurred image. Therefore, they are more efficient than the statistical priors and more accurate kernel are estimated.

Finally, we give one of our applications in Fig. 8, where two different focused images are fused to generate their all-in-focus images. As shown in Fig. 8, there is always one man out-of-focus in the input image (a) and (b). Using our method, we recover each blurred man in one image by taking the other image as the corresponding similar clear image. After our deblurring algorithm, both the men are focused in the restored images (c) and (d).

4. CONCLUSION

We have presented a new method to deblur an out-of-focus image from its similar clear images. With the constraints extracted from the similar images, we accurately estimated the blur kernel and produced high-quality results. By comparing with state-of-the-art methods, we demonstrated that similar clear images could provide useful information and improve the deblurring results. We just provided one way to efficiently utilize other similar images and more powerful algorithms will be developed in the future.

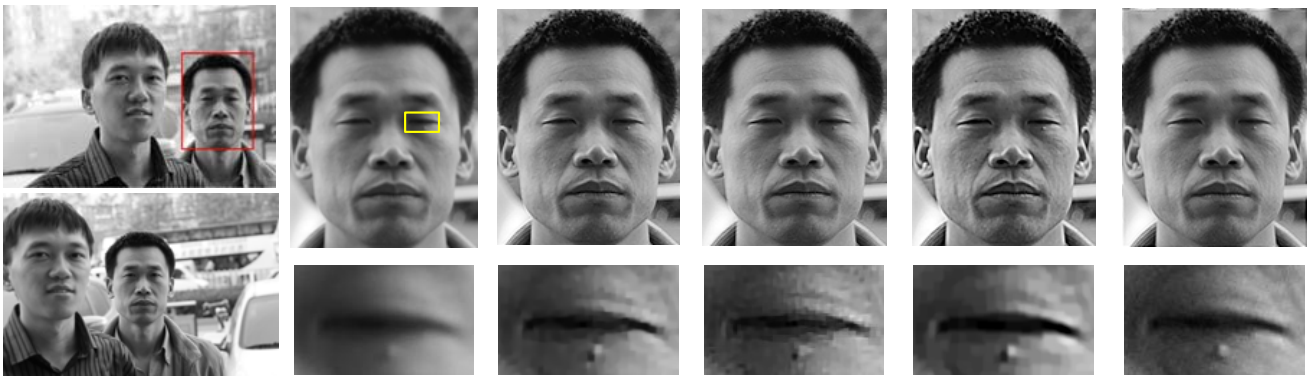

Fig. 6. Comparisons on real images. (a) Two input similar image pairs, the red rectangle is the region to be deblurred, (b) blurred region, (c)-(f) deblurred regions of [17], [3], [5] and ours respectively. The bottom row shows the close-ups of the yellow rectangle.

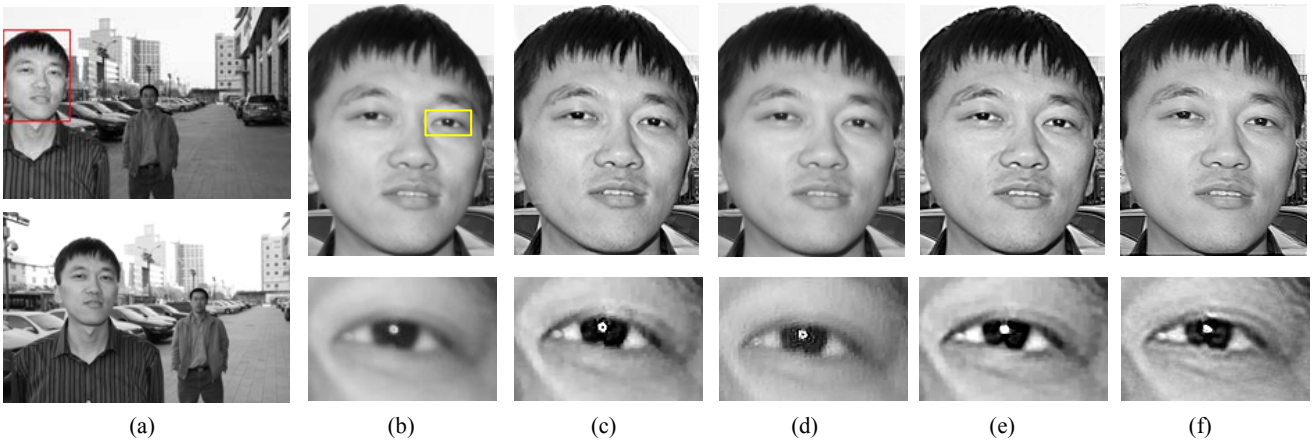

Fig. 7. Comparisons on real images. (a) Two input similar image pairs, the red rectangle is the region to be deblurred, (b) blurred region, (c)-(f) deblurred regions of [17], [3], [5] and ours respectively. The bottom row shows the close-ups of the yellow rectangle.

Fig. 8. Generating all focused images: (a) and (b) Two different focused images. The men in the green rectangles are in-focus and the men in the red rectangles are out-of-focus. The red rectangles are the regions to be deblurred. (c) the restored clear image of (a), (d) the restored clear image of (b). By taking one image as the blurred image and another image as the similar clear image, our method removes the blur in both input images. As shown, both (c) and (d) have the two men focused.

5. REFERENCES

[1] A. Levin, "Blind motion deblurring using image statistics," in *Proc. Neural Information Processing Systems*, British Columbia, Canada, Dec. 2006, pp. 841–848.

[2] R. Fergus and B. Singh, "Removing camera shake from a single photograph," *ACM Transactions on Graphics*, vol. 25, no. 3, pp. 787–794, 2006.

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

[4] D. Krishnan and R. Fergus, "Fast image deconvolution using hyper-laplacian priors," in *Proc. Neural Information Processing Systems*, British Columbia, Canada, Dec. 2009, pp. 1033–1041.

[5] D. Krishnan, T. Tay, and R. Fergus, "Blind deconvolution using a normalized sparsity measure," in *Proc. Intl Conf. Computer Vision and Pattern Recognition*, Colorado Springs, USA, June 2011, pp. 233–240.

[6] T. S. Cho, C. L. Zitnick, N. Joshi, S. B. Kang, R.Szeliski, and W. T. Freeman, "Image restoration by matching gradient distributions," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 34, no. 4, pp. 683–694, 2012.

[7] A. Rav-acha and S. Peleg, "Two motion-blurred images are better than one," *Pattern Recognition Letters*, vol. 26, no. 3, pp. 311–317, 2005.

[8] L. Yuan, J.n Sun, L. Quan, and H. Shum, "Image deblurring with blurred/noisy image pairs," *ACM Transactions on Graphics*, vol. 26, no. 3, pp. 1–10, 2007.

[9] S. Zhuo, D. Guo, and T. Sim, "Robust flash deblurring," in *Proc. Intl Conf. Computer Vision and Pattern Recognition*, San Francisco, USA, June 2010, pp. 2440–2447.

[10] S. Xiang, G. Meng, Y. Wang, C. Pan, and C. Zhang, "Image deblurring with matrix regression and gradient evolution," *Pattern Recognition*, vol. 45, no. 6, pp. 2164–2179, 2012.

[11] A. Levin, R. Fergus, F. Durand, and W. T. Freeman, "Image and depth from a conventional camera with a coded aperture," *ACM Transactions on Graphics*, vol. 26, no. 3, 2007.

[12] N. Joshi, S. B. Kang, C. L. Zitnick, and R. Szeliski, "Image deblurring using inertial measurement sensors," in *Proc. ACM SIGGRAPH*, Los Angeles, USA, July 2010, pp. 1–8.

[13] Ashutosh Saxena, Sung H. Chung, and Andrew Y. Ng, "3-d depth reconstruction from a single still image," *International Journal of Computer Vision*, vol. 76, no. 1, pp. 53–69, 2008.

[14] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[15] A. P. Pentland, "A new sense for depth of field," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 9, no. 4, pp. 522–531, 1987.

[16] M.A. Fischler and R.C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.

[17] L. Xu and J. Jia, "Two-phase kernel estimation for robust motion deblurring," in *Proc. European Conference on Computer Vision*, Crete, Greece, Sept. 2010, pp. 157–170.

[18] L. Lucy, "An iterative technique for the rectification of observed distributions," *Astronomical Journal*, vol. 79, no. 6, pp. 745–754, 1974.