SINGLE IMAGE BLIND DEBLURRING WITH IMAGE DECOMPOSITION

*Yuquan Xu, Xiyuan Hu, Lu Wang, Silong Peng**

Institute of Automation, Chinese Academy of Sciences, Beijing, 100190, P. R. China

ABSTRACT

How to deal with the motion blurred image is a common problem in our daily life. Restoring blurred images is challenging, especially when both the blur kernel and the sharp image are unknown. In this work, we present a new algorithm for removing motion blur from a single image, which incorporates the image decomposition into the image deblurring process. Most of the existing algorithms solving the blind deblurring problem use the alternate iterative mechanism, which alternative estimates the kernel and restores the sharp image. We find that the small gradients of image are not always helpful but sometimes harmful to this kind of iterative algorithm. So we decompose the blurred image into cartoon and texture components. And we only use the cartoon part of the image, which can improve the stability and robustness of the algorithm. Our experiments show that our algorithms can achieve good results in man-made and real-life photos.

Index Terms— Blind Deconvolution; Image Decomposition; Motion Deblurring; Image Enhancement;

1. INTRODUCTION

In our daily life, the blur caused by camera shake is very common and annoying. Recently image deblurring has received a lot of attention in the field of computer graphics and vision. In this work we aim to solving the uniform motion deblur problem, in which the blurred image can be modeled as a convolution of a sharp image and a blur kernel (also known as Point Spread Function, PSF). In this case the task of deblurring an image equals to image deconvolution.

Recently, Fergus et al.[1] have proposed a successful algorithm to restore blurry photos due to camera shake. In [1], Fergus et al. mentioned that they tried to search for the maximum a-posteriori (MAP) solutions , by alternately estimating the blur kernel and latent image gradient. However, this kind of manner failed to derive a satisfactory result. Thus, they used ensemble learning method [2] to solve the blind deblur problem, which is very complex and slow. Since then, many approaches have been proposed by introducing other constraints [3][4][5][6][7][8][9][10]. Yuan et al.[3] used blurry and noised image pair to restore the sharp image. Shan et al.[4] presented a local prior to detect and smooth surfaces. Joshi et al.[5] and Cho et al.[10] all adopted the edge information to estimate the kernel. Cho and Lee[6] applied bilateral filter and shock filter to predict salient edge. Levin et al. [7] analyzed and evaluated some blind deconvolution algorithms and presented that the naive $MAP_{x,k}$ prefer the kernel estimation as a delta function, but using MAP_k could get better result. Levin et al. [9] derived a simpler approximate MAP_k algorithm, and provided 5 different deconvolution algorithms. Compared with naive MAP method, most of the algorithms need complex optimization techniques that are slow and that do not always succeed. In this paper, we find that the image decomposition can play unexpected role in the image blind deconvolution. There are two reasons why the image separation can help solve image deblurring problem. In the beginning of the iteration, the kernel would not be estimated accurately, so there are many ringing and noise artifacts in the restored image. For removing these artifacts, we decompose the restored image and leave only the cartoon part. Secondly, the image decomposition can prevent the algorithm favor the no-blur result, for which the estimated kernel is a delta kernel and restored image is the input blurred image.

2. MAP FRAMEWORK

Shift-invariant motion blur is commonly modeled as a convolution process

$$\mathbf{B} = \mathbf{I} * \mathbf{k} + \mathbf{n} \tag{1}$$

where $\mathbf{B}$ is the observed blur image, $\mathbf{I}$ is the sharp image, $\mathbf{k}$ is the blur kernel, $\mathbf{n}$ is the noise, $*$ denotes convolution. By Bayes' theorem, we can write the posterior probablity:

$$p(\mathbf{I}, \mathbf{k}|\mathbf{B}) \propto p(\mathbf{B}|\mathbf{I}, \mathbf{k})p(\mathbf{I})p(\mathbf{k}) \tag{2}$$

where $p(\mathbf{B}|\mathbf{I}, \mathbf{k})$ represents the likelihood and $p(\mathbf{I})$ and $p(\mathbf{k})$ denote the priors on the latent image and the blur kernel. Since the image noise is modeled as Gaussian distribution, $p(\mathbf{B}|\mathbf{I}, \mathbf{k}) \propto \prod_i N(\mathbf{B}(i) - \mathbf{I} * \mathbf{k}(i)|0, \sigma^2)$. We choose Laplace prior as the image prior, which is written as $p(\mathbf{I}) \propto \prod_i \exp(\partial I(i))$. Since the motion kernel corresponds the path of the camera, it should be sparse, with most values

*Research supported in part by the National Natural Science Foundation of China (60972126,10801004), the Joint Funds of the National Natural Science Foundation of China (U0935002/L05) and the Beijing Natural Science Foundation (4102060).

Algorithm 1 The kernel estimation algorithm

1: **Required**: The blurred image B
2: **Required**: The initial kernel size and λ
3: **repeat**
4: Estimate the latent image update (7)
5: Decompose the latent image and keep the cartoon part using algorithm of [12]
6: Estimate the blur kernel update (8)
7: **until** meets the maximum iteration number
8: **return** Restored image and kernel.

close to zero. We therefore model the prior on blur kernel as exponentially distributed, $p(\mathbf{k}) \propto \prod_j \exp(-k_j), k_j > 0$. Finally the posterior distribution has the form:

$$\begin{aligned} p(\mathbf{I},\mathbf{k}|\mathbf{B}) &\propto p(\mathbf{B}|\mathbf{I},\mathbf{k})p(\mathbf{I})p(\mathbf{k}) \\ &\propto \prod_i N(\mathbf{B}(i) - \mathbf{I} * \mathbf{k}(i)|0,\sigma^2) \\ &\quad \prod_i \exp(\partial\mathbf{I}(i)) \prod_j \exp(-\mathbf{k_j}) \end{aligned} \tag{3}$$

where i indexes over image pixels and j indexes over blur kernel elements. And the estimated kernel and restored image are the solutions of the MAP problem, $\mathbf{I},\mathbf{k} = \arg\max_{\mathbf{I},\mathbf{k}} p(\mathbf{I},\mathbf{k}|\mathbf{B})$. Our MAP problem is easily transformed to an energy minimization problem with denoting $E(\mathbf{I},\mathbf{k}) = -\log(p(\mathbf{I},\mathbf{k}|\mathbf{B}))$. So the MAP problem is equivalent to:

$$\min_{\mathbf{I},\mathbf{k}} E(\mathbf{I},\mathbf{k}) = ||\mathbf{I} * \mathbf{k} - \mathbf{B}||_{\mathbf{2}}^{\mathbf{2}} + \lambda_{\mathbf{1}}||\nabla\mathbf{I}||_{\mathbf{1}} + \lambda_{\mathbf{2}}||\mathbf{k}||_{\mathbf{1}} \tag{4}$$

In order to optimize (4), the alternate iterative mechanism is used as iteratively estimating $\mathbf{I}$ and $\mathbf{k}$.

$$\mathbf{I} = \arg\min_{\mathbf{I}} ||\mathbf{I} * \mathbf{k} - \mathbf{B}||_2^2 + \lambda_1||\nabla\mathbf{I}||_1 \tag{5}$$

$$\mathbf{k} = \arg\min_{\mathbf{k}} ||\mathbf{I} * \mathbf{k} - \mathbf{B}||_2^2 + \lambda_2||\mathbf{k}||_1 \tag{6}$$

Unfortunately, the naive MAP solutions prefer the no-blur result. As in Figure 1 (b), the deblurring image with naive MAP is just the blurred input and kernel estimation is delta function. It seems like that the algorithm don't work at all.

3. OUR DEBLURRING ALGORITHM

Levin et al. [9] presented that using the image gradient instead of image value itself can significantly improve the deconvolution results. In Figure 1 (c), when we use the gradient of the image, the kernel estimation is not the delta function but still far away from the ground truth kernel. In practice we find that the small gradients of image is not always helpful but sometimes harmful to this kind of iterative algorithm, so we threshold the gradient like [6]. Figure 1 (d) shows the result. The kernel estimated by the threshold gradient have the general shape of a real kernel, but is still not very accurate. Finally, we find that together with image decomposition the MAP method works quite well, as shown in Figure 1 (e). The image decomposition approaches are to separate the image into two components: one represents a cartoon or piece-wise smooth component; and the other represents the oscillatory or texture component. The cartoon part of the image remove the texture of the image while keeping the strong edges. Here, we use an operator-based[12] approach to fulfill such kind of image separation, which model the cartoon part of the image as in the null space of some operators. There are two main reasons explained why using the cartoon part image can help the image deblurring. (1) in the beginning of the iteration, the kernel estimation will have a lot of errors, so there are many ringing and noise artifacts in the restored image. Using the cartoon part of the image can decrease these artifacts and help the algorithm. (2) using the cartoon part can further prevent the algorithm favor the no-blur result. Just like [8], we change (5) into a easy one (7), which can be solved very fast in frequency domain. Experiments show although (7) may not produce high-quality results but it don't hinder accurate estimation of kernel. The final cost functions for update latent image $\mathbf{I}$ and blue kernel $\mathbf{k}$ have following forms:

(a) (b) (c) (d) (e)

Fig. 1: (a)blurred image and real kernel; (b) result using naive MAP; (c) result using image gradient; (d)result using image gradient after threshold; (e) result using image gradient after threshold of cartoon part

$$\mathbf{I} = \arg\min_{\mathbf{I}} ||\mathbf{B} - \mathbf{I} * \mathbf{k}||_2^2 + \lambda_1||\nabla\mathbf{I} - \nabla\mathbf{I}_{car,thr}||_2^2 \tag{7}$$

$$\mathbf{k} = \arg\min_{\mathbf{k}} ||\nabla\mathbf{B} - \nabla\mathbf{I}_{car,thr} * \mathbf{k}||_2^2 + \lambda_2||\mathbf{k}||_1 \tag{8}$$

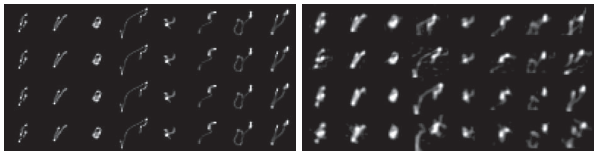
Where $\nabla\mathbf{I}_{car,thr}$ means the gradients after threshold of the cartoon image. In the experiment, we set the $\lambda_1 = 0.02$ and $\lambda_2 = 1$ for all images. And we estimate the blur kernel and latent image in a multi-scale setting, which is proved to be an effective scheme in blind deblurring problem. The final algorithm is show in Algorithm 1

4. EXPERIMENTS AND RESULTS

In this section, we present experiment results of our algorithm. We compare it to the methods in [6][9], the implements of these two paper can be downloaded from internet . In order to compare the results of kernel estimation with more accurate, we use the same deconvolution method in [13] with the same

Table 1: Comparison of the SSD of 3 algorithms

		ker1	ker2	ker3	ker4	ker5	ker6	ker7	ker8	average
image1	Cho and Lee[6]	**33.50**	**35.33**	**30.65**	433.80	**27.04**	51.59	102.62	240.49	119.37
	Levin et al.[9]	44.42	48.89	35.24	97.11	34.02	28.66	46.85	48.68	47.98
	Our method	39.55	48.00	31.19	**71.33**	37.88	**24.95**	**34.78**	**39.31**	**40.87**
image2	Cho and Lee[6]	70.81	**43.71**	**32.06**	111.91	**30.50**	95.06	97.11	115.80	74.62
	Levin et al.[9]	64.41	74.53	64.61	**105.92**	36.15	36.95	64.45	71.69	64.84
	Our method	**55.25**	63.25	47.70	113.38	35.91	**29.66**	**52.62**	**46.27**	**55.50**
image3	Cho and Lee[6]	**25.11**	**33.40**	**17.87**	45.04	27.33	53.19	52.85	41.67	37.06
	Levin et al.[9]	39.06	47.14	19.19	68.81	17.60	21.16	27.24	31.53	33.96
	Our method	32.36	40.05	21.06	**41.82**	**15.10**	**14.43**	**16.55**	**26.17**	**25.94**
image4	Cho and Lee[6]	**37.04**	**75.12**	**19.59**	254.38	**21.84**	70.33	**42.12**	79.81	75.03
	Levin et al.[9]	78.38	129.02	50.08	98.03	34.08	62.52	98.19	112.99	82.91
	Our method	47.49	90.10	28.26	**69.13**	35.01	**29.31**	58.29	**37.12**	**49.34**



(a) Groundtruth kernel (b) Cho and Lee [6]

(c) Levin et al. [9] (d) Our method

Fig. 2: Recovered kernels, for the set of 32 test images

parameter settings, so the deconvolution results only depend on estimated kernels.

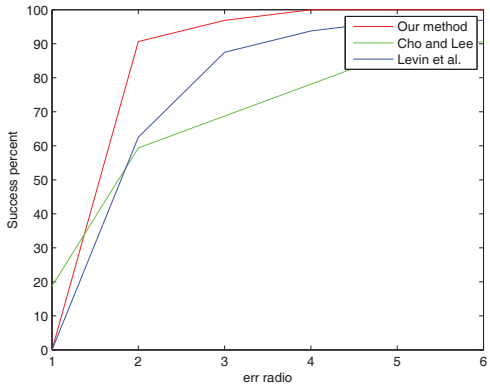


Fig. 3: Cumulative error ratios for the algorithms, Cho and Lee [6], Levin et al. [9] and our algorithm

We first test the algorithm on synthetic examples from the dataset in [7]. In this dataset, four 255×255 gray-scale images are blurred by 8 different kernels which lead to 32 test images. The blurred data, ground truth data and ground truth kernels are provided. We compare our kernel estimation results with the blind deconvolution algorithms of [6][9]. Because [9] provide five different deconvolution methods, we choose the "sparse, free-eng, filt space" one to be compared, which has the best performance in [9].

The results of the all the kernel estimation with 3 algorithm are showed in Figure 2. Figure 2 shows the kernel estimations by our algorithm are closest to the ground truth, especially in the fourth row of the kernel estimations. To better illustrate this point, we show all the SSD (sum of squared differences) errors on 32 test images using these 3 methods in Table 1. For different images and different kernels, the performance of the algorithms is changed. In all 32 test images, Cho and Lee have 15 best results, Levin et al. only have 1, and our algorithm have 16. It is noteworthy that although Cho and Lee have best resuls in almost half of images, it has the biggest average error, which means Cho's algorithm can deblur some images very well but also fails dramatically in some other examples. From Table 1, we can see that our method gets better performance in almost all the test images than algorithm proposed in [9], and has the smallest error in the average sense, which means the robustness of our algorithm is the best. When the Cho and Lee [6] fails as in Figure 4, our algorithm is still able to recover a reasonable kernel.

We compute the ratios of SSD error (1) between ground truth image and deconvolved image with estimated kernel; (2) between ground truth image and deconvolved image with true kernel. The error metric used is the same in [7][9]. In Figure 3, we show the cumulative error ratios for these 3 algorithms. We can easily see that our algorithm has the best performance. As mentioned in [7], when the error ratios are above 2, the restored image are already visually implausible. The error radio in Figure 3 show that 90.6% of our results are below 2 compared with 59.4% of Cho and 62.5% of Levin. From a global point of view, our algorithm brings a satisfying solution to these test images, and gets the best performance.

(a) SSD=48.34 (b) SSD=433.80 (c) SSD=71.33

Fig. 4: Synthetic image results for image 1 and kernel 4 in the dataset. (a) is the deblurring result using the ground truth kernel; (b) is the result with kernel estimated by Cho and Lee [6]; (c) is our result

Then we compare the deblurring results of these three algorithms on some real captured image presented in some other blind deconvolution papers whose real kernels are unknown. Figure 5 shows the deblurred results. (f) shows the word in the top of the building is not clear, and more rings and artifacts around the distant buildings are shown in (b) and (c). Overall our algorithm can give shaper restored image and has less artifacts than other algorithms.

5. REFERENCES

[1] R. Fergus, B. Singh, A. Hertzmann, S.T. Roweis, and W.T. Freeman, "Removing camera shake from a single photograph," *ACM Trans. Graph.*, vol. 25, pp. 787–794, 2006.

[2] J. Miskin and D.J.C. MacKay, "Ensemble learning for blind image separation and deconvolution," in *Adv. in Independent Component Analysis*, 2000.

[3] L. Yuan, J. Sun, L. Quan, and H. Shum, "Image deblurring with blurred/noisy image pairs," *ACM Trans. Graph.*, vol. 26, 2007.

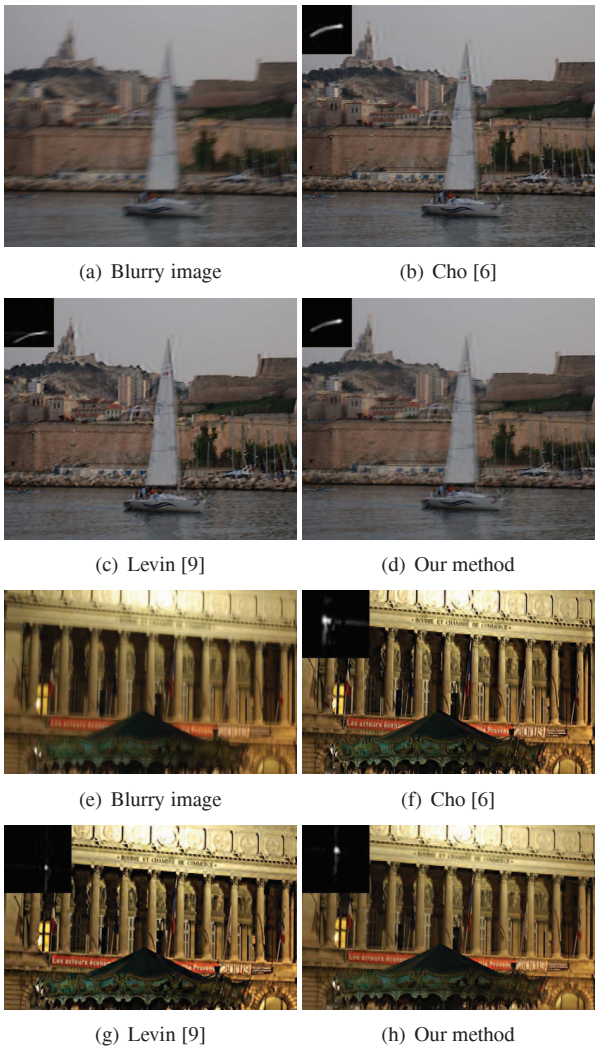
[4] Q. Shan, J. Jia, and A. Agarwala, "High-quality motion deblurring from a single image," *ACM Trans. Graph.*, vol. 27, no. 3, pp. 73:1–73:10, 2008.

[5] N. Joshi, R. Szeliski, and D.J. Kriegman, "Psf estimation using sharp edge prediction," in *CVPR*, 2008, pp. 1–8.

[6] S. Cho and S. Lee, "Fast motion deblurring," *ACM Trans. Graph.*, vol. 28, pp. 145:1–145:8, 2009.

[7] A. Levin, Y. Weiss, F. Durand, and W.T. Freeman, "Understanding and evaluating blind deconvolution algorithms," in *CVPR*, 2009, pp. 1964 –1971.

[8] L. Xu and J. Jia, "Two-phase kernel estimation for robust motion deblurring," in *ECCV*, 2010, ECCV'10, pp. 157–170.



(a) Blurry image (b) Cho [6]

(c) Levin [9] (d) Our method

(e) Blurry image (f) Cho [6]

(g) Levin [9] (h) Our method

Fig. 5: Real image results

[9] A. Levin, Y. Weiss, F. Durand, and W.T. Freeman, "Efficient marginal likelihood optimization in blind deconvolution," in *CVPR*, 2011, pp. 2657 – 2664.

[10] T.S. Cho, S. Paris, B.K.P. Horn, and W.T. Freeman, "Blur kernel estimation using the radon transform," in *CVPR*, 2011, pp. 241 – 248.

[11] D. Krishnan, T. Tay, and R. Fergus, "Blind deconvolution using a normalized sparsity measure," in *CVPR*, 2011, pp. 233 – 240.

[12] Xiyuan Hu, Weiping Xia, Silong Peng, and Wen-Liang Hwang, "Multiple component predictive coding framework of still images," in *ICME*, july 2011, pp. 1 –6.

[13] A. Levin, R. Fergus, F. Durand, and W.T. Freeman, "Image and depth from a conventional camera with a coded aperture," *ACM Trans. Graph.*, vol. 26, 2007.