Two-step Facial Images Deblurring With Kernel Refinement Via Smooth Priors

Jun-Hong Chen[1] and Long-Wen Chang[2]
Department of Computer Science
National Tsing Hua University
Hsinchu,Taiwan
Email[1]: jhchen.nthu@gmail.com
Email[2]: lchang@cs.nthu.edu.tw

Abstract— **Image deblurring is a challenging task in image processing. It is an ill-posed problem to estimate the unknown blur kernel and recover the original image from a blurred image. There are many methods for blurred natural images; however, few of them are able to perform well on blurred face images. Based on L_0 norm prior, we propose a two-step method for the images deblurring. The proposed method does not require any facial dataset to initialize the gradient of contours or any complex filtering strategies. In the first step, we combine L_0 norm prior with our local smooth prior to predict the blur kernel. With simple Gaussian filtering, we could maintain the smooth region in the latent image. In the second step, we refine the previous estimated kernel. In order to discard low intensity pixels that seemed to be noises on the kernel, we impose the sparsity on the kernel. Experimental results demonstrate that our proposed algorithm performs well on the facial images.**

Keywords-image debluring; kernel; smooth prior.

I. INTRODUCTION

Blind deblurring is a tough task in image processing. It can be modelled as

$$B = x * k + n, \quad (1)$$

where B is an observed blurred image, x is a latent image, $*$ is a convolution operator, k is a blur kernel, and n is a noise. Recovering a sharp image from the single blurred image B is an ill-posed problem. It can be solved by infinite set of pairs of the blur kernel and the latent image. Recent researches have a significant advance on image deblurring by iterative methods to estimate x and k. We can obtain a recovered image and an estimated kernel in each iteration. Usually, the temporarily recovered image is referred to as a latent image. The latent image is a clear and deblurred image. With edge information, the latent image and the blur kernel are estimated iteratively [3][5][8][10]. Shan et al. [2] adopted a sparse image prior via global and local priors. Krishnan et al. [6] used the L_1/L_2 prior on the high frequencies of an image. Xu et al. [9] used an unnatural L_0 prior to select high gradient edges iteratively. Kotera et al. [11] used straightforward maximum a posteriori method with $L_{0.3}$ heavy-tail prior. In this paper, we add a local prior to maintain smooth regions and refine the estimated kernel by the L_0 prior. In contrast to Pan et al. [14], which is focused on facial image deblurring, we predict the latent image without dataset supports and our experiments are better in image quality with appropriate parameters. In Section 2, related works are addressed. In Section 3, our proposed method is introduced. Section 4 describes our experimental results and Section 5 is the conclusion.

II. RELATED WORK

Image deblurring can be divided into two parts: blind deblurring and non-blind deblurring. The former is to recover a sharp image with only a given blurred image; the latter uses deconvolution to get a better result than the former. The challenge on facial images deblurring is that there is less texture in facial images. The existing methods use few edges in the blurred image to estimate a blur kernel. Cho and Lee [3] used an explicit filter as the shock filter to process the blurred image to select its real edges. Xu et al. [5] created a metric to measure the usefulness edges which is defined by local gradient information. Bae et al. [8] focused on the informative edges in patches. They combined gradient magnitude, the edges of straightness and usefulness edges as mentioned above. Other works have constraints on the sparsity of image gradients [2][6][9][11]. Shock filter has been used in [3] as a sharpened reference image for deblurring. However, due to severe blurring motion, recovering facial contour is difficult due to wrong edge selection. The gradient of heavy-tail distribution [1][2][4][7][11] property on face images might not work well because of less textures in those images. In other works, different sparsity priors [6], such as the normalized prior $\frac{\|x\|_1}{\|x\|_2}$ were introduced. According to Xu et al. [9], we can use it to recover the latent image with only several iterations. Without extra filtering such as shock filters, the optimization progress is faster than other methods. Recently, Pan et al. [14] proposed a deblurring method which uses similar face contours in their exemplar dataset for initial guess, and solve the objective function by the L_0 prior of the gradient magnitude. They collected hundreds of images and the gradients of those images to build their exemplar dataset. The success of the method is due to the facial global structure which has similar contours. However, an additional dataset is required to choose the initial guess. Different from the method of Pan et al. [14], we propose a two-step method to refine our results without dataset support. We focus on the smooth regions in face images and preserve the flat regions

 ISBN: 978-1-61208-602-6

by a local smooth term to alleviate the problem of ambiguous edge selection.

III. PROPOSED METHOD

The proposed method is shown in Figure 1. The blurred image is processed with the kernel estimation to obtain the initial kernel. Then, with the adjustment of the initial kernel based on its sparsity, we can get the final estimated kernel and the deblurred image. The kernel estimation is done with a coarse-to-fine approach. We estimate the latent image (x step) and the blur kernel (k step) iteratively in different scales. The coarse-to-fine method solves x by minimizing the objective function with a local smooth term from a low image scale to a high image scale. The estimated kernel at the last scale will be the initial kernel next round. In the kernel adjustment, the predicted kernel is denoised by using L_0 regularization to discard its low intensity pixels. The objective function with our local smoothness term is

$$\min_{x,z}\|x * k - B\|_2^2 + \lambda\|\nabla x\|_0 + \alpha\|\nabla x - \nabla z\|_2^2 \quad (2)$$

where x is a latent image, B is a blurred image, and k is a blur kernel, ∇ is a gradient operator of two directions (horizontal and vertical).

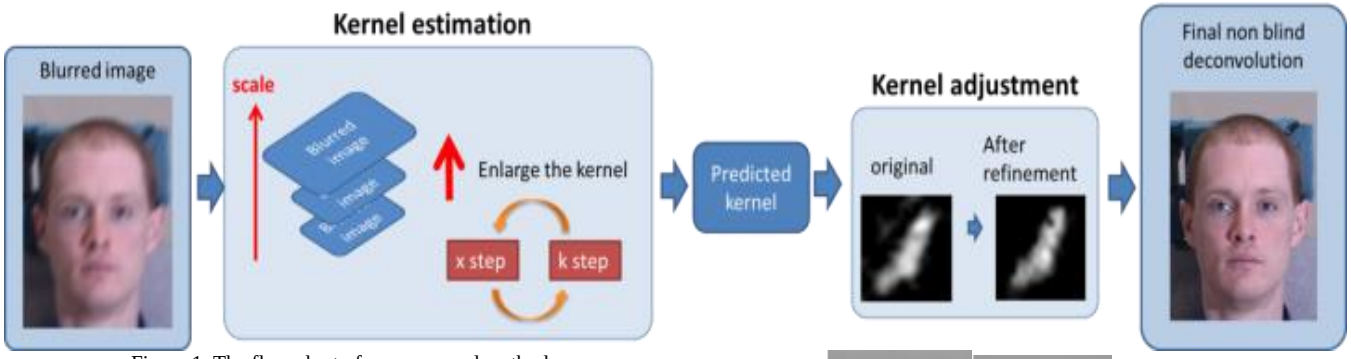


Figure 1. The flow chart of our proposed method

The first term represents the error between the observed blurred image and the estimated blurred image convoluted by the latent image and the kernel. The second term represents the constraint of the sparsity in the latent image. The third term is our proposed local smoothness term that we want the latent image to be close to the z map which indicate the smooth regions in the blurred image. As in Pan's [14], we can rewrite (2) as

$$\min_{x,w,z}\|x * k - B\|_2^2 + \beta\|\nabla x - w\|_2^2 + \lambda\|w\|_0 + \alpha\|\nabla x - \nabla z\|_2^2, \quad (3)$$

and

$$w = \begin{cases} \nabla x, & if\ \ |\nabla x|^2 \geq \frac{\lambda}{\beta}, \\ 0, & otherwise. \end{cases} \quad (4)$$

The variable w helps us to solve L_0 term. The larger β becomes, the closer the solution of (3) approaches (2). Because of the multiple-variable minimization of (3), we solve it alternatively by updating w, z and x independently. Given the latent image x, we obtain w by (4) with a threshold. Here, λ is a constant, and β becomes larger in each iteration by a factor of 2. In this function, we only choose high gradient pixels in the blurred image which makes the latent image sparse. We set

$$z = \sim t \circ \hat{x}, \quad (5)$$

$\hat{x}$ is the latent image, which is filtered by a Gaussian filter with initial variance $\sigma = 0.3$ and increases over each scale, $\circ$ is a pointwise product (pixel to pixel), and ~ is a NOT operator. t is a binary image indicating high gradients in the latent image (using w),

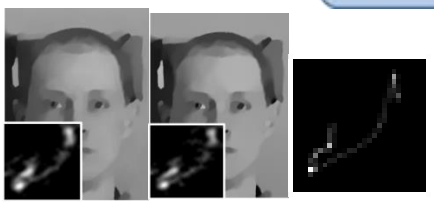
$$t = \begin{cases} 1, & w \neq 0 \\ 0, & w = 0' \end{cases}$$

We dilate t to avoid the influence on high gradient edge,

$$t = t \oplus m,$$

where $\oplus$ is a dilation operator, m is a square structuring element whose width is a half of estimated kernel size.

In our experiment, we set the initial kernel size for 3 pixel, and increases it over each scale by factor of $\sqrt{2}$. Thus, the z map we introduced could be thought of as the set of all pixels in the latent image with filtering, which is exclusive of the high gradient pixels in each iteration. According to Shan et al. [2], the smooth region in the latent image is still smooth after motion blur. The results with our smooth term are shown



(a) (b) (c)

Figure 2. The latent image and estimated kernel recovered by our method with and without our priors of local smoothness. (a) Our local smooth prior ($\alpha = 0.0008$, $\lambda = 0.02$) (b) Without our local smooth prior ($\alpha = 0$, $\lambda = 0.02$) (c) original kernel

in Figure 2 (a). Minimizing (6) with respect to x, we obtain

$$\min_x\|x * k - B\|_2^2 + \beta\|\nabla x - w\|_2^2 + \alpha\|\nabla x - \nabla z\|_2^2 . \quad (6)$$

According to Pan et al. [14], we can get

$$x = \mathcal{F}^{-1}\left(\frac{\overline{\mathcal{F}(k)}\mathcal{F}(B) + \beta\left(\overline{\mathcal{F}(\partial_x)}\mathcal{F}(w_x) + \overline{\mathcal{F}(\partial_y)}\mathcal{F}(w_y)\right) + \alpha\left(\overline{\mathcal{F}(\partial_x)}\mathcal{F}(z_x) + \overline{\mathcal{F}(\partial_y)}\mathcal{F}(z_y)\right)}{\overline{\mathcal{F}(k)}\mathcal{F}(k) + (\beta+\alpha)\left(\overline{\mathcal{F}(\partial_x)}\mathcal{F}(\partial_x) + \overline{\mathcal{F}(\partial_y)}\mathcal{F}(\partial_y)\right)}\right) \quad (7)$$

where $\mathcal{F}^{-1}(\cdot)$ and $\mathcal{F}(\cdot)$ are the Inverse Discrete Fourier Transform (IDFT) and Discrete Fourier Transform (DFT), and ∂_x and ∂_y are derivative operators in x and y directions, and $\bar{\cdot}$ is the complex conjugate operator. The objective function of the kernel estimation is

$$\min_k \|x * k - B\|_2^2 + \gamma\|k\|_2^2 \quad (8)$$

where x is a latent image, B is a blurred image, and k is a blur kernel, γ is the parameter for regularization term on the kernel sparity. It can be solved by the conjugate gradient decent easily. Here, the data term is on gradient level according to Pan et al. [14], which has a stable kernel estimation can be written as

$$\min_k \|\nabla x * k - \nabla B\|_2^2 + \gamma\|k\|_2^2, \quad (9)$$

where ∇ is gradient of horizontal and vertical direction. It is efficient to rewrite (9) in the matrix form. Thus, (9) can be rewritten as below:

$$(\nabla x\mathrm{K} - \nabla B)^{\mathrm{T}}(\nabla x\mathrm{K} - \nabla B) + \gamma\mathrm{K}^{\mathrm{T}}\mathrm{K} = \mathrm{K}^T\nabla x^T\nabla x\mathrm{K} - \mathrm{K}^T\nabla x^T\nabla B - \nabla B^T\nabla x\mathrm{K} + \gamma\mathrm{K}^{\mathrm{T}}\mathrm{K}, \quad (1)$$

where K is a convolution matrix referred to k. By the minimization of (10) with respect to K we can obtain

$$(\nabla x^T\nabla x + \gamma)\mathrm{K} = \nabla x^T\nabla B \quad (11)$$

Then, (11) can be solved by the conjugate gradient decent. The coarse-to-fine kernel estimation is with (11) and estimation of the latent image is with (4), (5) and (7). The experimental observation of output kernel shows that there are still some low intensity pixels that seemed to be noise. The straightforward concept is using L_0 norm to let the kernel be sparse. Based on L_0 norm method, this problem involves the following terms

$$\min_k \|\nabla x * k - \nabla B\|_2^2 + \lambda\|k\|_0, \quad (12)$$

Which represents the constraint on the sparsity of blur kernel. Since the function in equation (12) is a non-convex function, we add an auxiliary variable r into (12) and rewrite it as the previous case

$$\min_{k,r} \|\nabla x * k - \nabla B\|_2^2 + \lambda\|r\|_0 + \beta\|k - r\|_2^2. \quad (13)$$

In each iteration, we alternatively solve subproblems with respect to each variable r and k,

$$r = \begin{cases} k, & if\ |k|^2 \geq \frac{\lambda}{\beta} \\ 0, & otherwise \end{cases}. \quad (14)$$

Here, λ is a constant, and β increases by a factor of 2 in each iteration. The threshold of the intensity of kernel pixels is to discard the low value of pixels.

$$\min_k \|\nabla x * k - \nabla B\|_2^2 + \beta\|k - r\|_2^2. \quad (15)$$

After the variable r is obtained, minimization of (14) can be rewritten as below:

$$(\nabla x^T\nabla x + \beta)\mathrm{K} = \nabla x^T\nabla B + \beta\mathrm{R}, \quad (16)$$

where K is a convolution matrix referred to k, R is a convolution matrix referred to r.

We can solve (16) by conjugate gradient decent. The concept in the second step is based on the progressive sparsity of L_0 norm. Because L_0 norm is a non-convex function, we have to use additional variables and iterations to solve. Therefore, in the k step (kernel estimation process), we estimate the kernel with L_2 norm regularization which can be done easily by conjugate gradient decent. Compared to Pan et al. [14], we rewrite the L_0 method in coarse-to-fine approach. Without dataset support, we solve it in a single image. Compared to Xu et al. [9], the local smooth prior is introduced, and we use two steps to refine our predicted kernel. Figure 2 shows the refinement of kernel estimation. We can see that Figure 3 (b) shows the kernel after refinement. It is much closer to the kernel of ground truth than Figure 2(a) the kernel before the refinement.

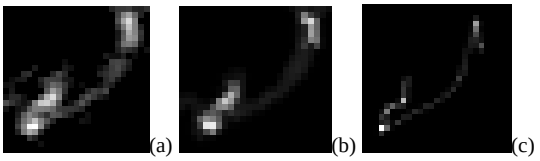

(a) (b) (c)

Figure 3. Refinement of kernel estimation (a) Before (b) After (c) Ground truth kernel

Finally, we obtain the refinement kernel and use the non-blind deconvolution provided by Pan et al. [14]: *deconvSps* function to get the recovering image.

IV.EXPERIMENTS

In our experiments, the parameters: α = [4e−4 ~ 64e−4], λ=[0.02,0.03,0.04], and initial kernel size is 17×17. The dataset we test is provided by Pan et al. [14] and Levin et al. [11]. The ground truth images are the deconvolution of the blur images with the ground truth blur kernel. The non-blind deconvolution method is provided by Pan et al. [14]: deconvSps function, which is the same method we used in our estimated kernels to recover the sharp images for consistent measures. To validate our framework, we compare the visual quality of the recovered sharp latent image and kernel which is shown in Figure 4 without our local smooth prior and Figure 5 with and without our kernel refinement. Figure 6 shows the comparison of the results by various methods. The recovering method, deconvSps function, is the same provided by Pan et al. [14]. The image quality shows that the image recovered by our kernel looks softer and smoother. In addition, our local smooth prior and kernel adjustment could reduce the noise of the estimated kernel; that is, it is reasonable to fit two-step approach to improve the kernel. Also, 40 blur images provided by Pan et al. [14] are used. PSNR (Peak Signal to Noise Ratio) and SSIM (Structural Similarity) are the metrics between the images with ground truth kernel via non blind deconvolution and the deblurred images. Figures 7 and Figure 8 show that the comparison of PSNR and SSIM values for 40 images and our method has better performance.

IV. CONCLUSION

We present a new framework combined with L_0 norm prior for images which leverages those smooth regions and refines the kernel to get better results. The local smooth term is to maintain the smoothness in images. Without any dataset support, we use coarse-to-fine approach and perform well on face images.The better results could be credited to the flat region in these face images to alleviate the ambiguous edge selection.

 ISBN: 978-1-61208-602-6

REFERENCES

[1] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman, "Removing camera shake from a single photograph," ACM Transaction on Graphics, 25(3):787–794, 2006.

[2] Q. Shan, J. Jia, and A. Agarwala, "High-quality motion deblurring from a single image," ACM Transaction on Graphics, 27(3):73;1-73:10, 2008.

[3] S. Cho and S. Lee, "Fast motion deblurring," ACM Transaction on Graphics, 28(5), 145:1-145:8,2009.

[4] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman, "Understanding blind de-convolution algorithms," IEEE Transaction on Pattern Analysis AND Machine Intelligence, Vol. 33, No. 12, 2354-2367, 2011

[5] L. Xu and J. Jia: "Two-phase kernel estimation for robust motion deblurring," Proc. 11th European Conference on Computer Vision, 2010

[6] D. Krishnan, T. Tay, and R. Fergus, "Blind deconvolution using a normalized sparsity measure,". Computer Vision and Pattern Recognition, 2011.

[7] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman, "Efficient marginal likelihood optimization in blind deconvolution, " Proc. IEEE Conf Computer Vision and Pattern Recognition, 2011.

[8] H. Bae, C. C. Fowlkes, and P. H. Chou, "Patch mosaic for fast motion deblurring," Proc. 11th Asian Conference on Computer Vision, 2012.

[9] L. Xu, S. Zheng, and J. Jia, "Unnatural l0 sparse representation for natural image deblurring," Proc. IEEE Conf Computer Vision and Pattern Recognition, 2013.

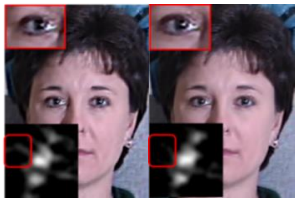
[10] L. Sun, S. Cho, J. Wang, and J. Hays, "Edge-based blur kernel estimation using patch priors," Proc. IEEE International Conference on Computational Photography, 2013

[11] J. Kotera, F. Šroubek, and P. Milanfar, "Blind deconvolution using alternating maximum a posteriori estimation with heavy-tailed priors," 15th International Conference on Computer Analysis of Images and Patterns, 2013.

[12] T. Michaeli and M. Irani, "Blind Deblurring Using Internal Patch Recurrence," Proc 13th European Conference European Conference on Computer Vision, 2014

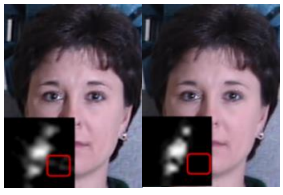
[13] J. Pan, Z. Hu, Z. Su, and M.-H. Yang, "Deblurring Text Images via L_0-Regularized Intensity and Gradient Prior," Proc. IEEE Conf Computer Vision and Pattern Recognition,, 2014.

[14] J. Pan, Z. Hu, Z. Su, and M.-H. Yang, "Deblurring face images with exemplars." Proc 13th European Conference European Conference on Computer Vision , 2014.



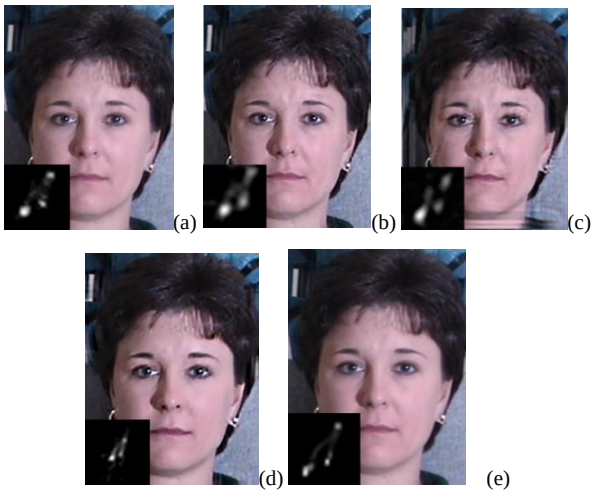
(a) (b)

Figure 4 (a) without our local smooth prior. α=0, λ=0.02 (b) With our local smooth prior. α=16e-4, λ=0.02



(a) (b)

Figure 5 (a) Without our kernel refinement (b) With our kernel refinement



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

Figure 6. Comparison of Image quality (a) Our method (b) Pan [14] (c) Kotera [11] (d) Xu [5] (e) Ground truth kernel with non-blind deconvolution

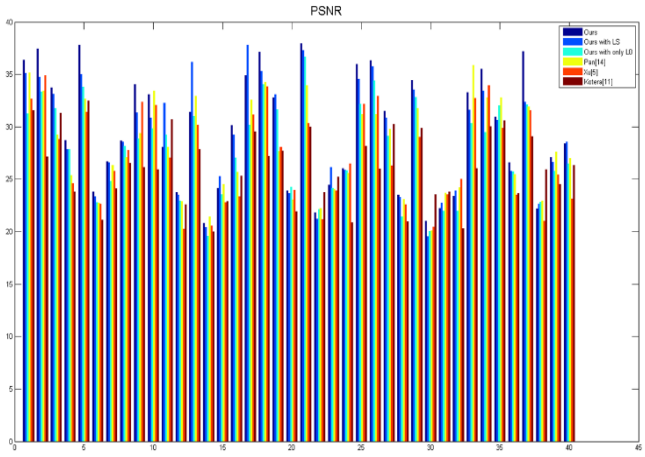


Figure 7. Comparison of PSNR values for 40 blurred images

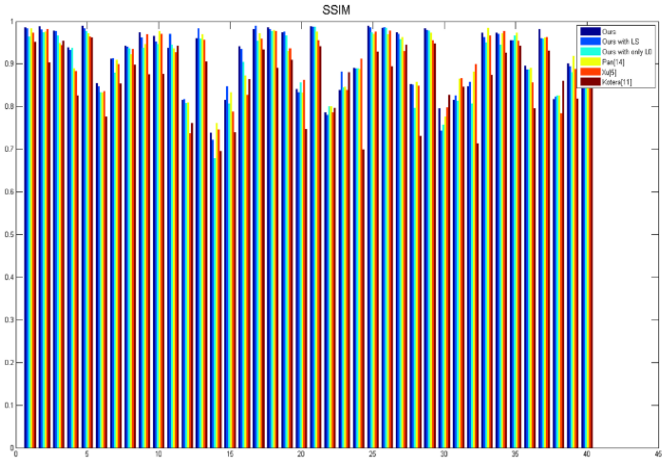


Figure 8. . Comparison of SSIM values for 40 blurred images