

A BLIND DECONVOLUTION APPROACH TO SHAKEN IMAGE

Quang Thi Nguyen, Jun Xi Sun and Yang Sun

School of Electronics and Information Engineering, Changchun University of Science and Technology, 7089 Weixing Road, Changchun, Jilin 130022, P. R. China

ABSTRACT

Blurry image are the bane of many photographers. The blur, which the camera moved during the exposure, is a common phenomena of image degradation. This movement may be very small but still creates blurry images. In this paper, a new method is proposed for deblurring the blurry image by applying a multi-scale blind-deblurring approach, which based on variational Bayesian estimation theory and explicitly handles outlier. In this method, the kernel estimation operation is behaved by varying image resolution in a coarse-to-fine technique, and then the component for outlier modeling is exploited in deconvolution process. In addition, an Average Difference criterion is incorporated for evaluating image quality. Starting with an image that has been blurred by camera shake, the unknown shape image is recovered in two phases: (1) a PSF estimation phase using variational Bayesian method, and (2) the shape image recovery by refining the outlier classification and the latent image based on Expectation-Maximization (EM) algorithm. The experimental approach is shown that the restored images are better than those given by the algorithm in other methods from previous works.

KEYWORDS: Camera Shake; Blind Deconvolution; Blur Kernel; Kernel Estimation

I. INTRODUCTION

Camera shake is a common effect in many images, since the resulting blur spoils many photos taken in low-light conditions. The shake effect can be reduced by using faster exposures, but cannot avoid other effects such as sensor noise or ringing effect. Much significant progress has been made recently towards removing this blur from images. The camera shake effect can be modeled as blur kernel in the convolution with latent image, and recovering unknown latent image is one of the toughest challenges in image processing. There have been many blind deconvolution methods described in recent literature. Many of these methods haven't considered the outliers, such as pixel saturation, non-Gaussian noise and nonlinear camera response curve that cause severe ringing artifacts in final output images. Q. Shan et al. [14] proposed a natural image statistics combine with L2-norm based data fidelity term, that can obtain accuracy results but this method only behaves in case the observed image contain a small of noise. Harmeling et al. [9] proposed a method for handling saturated pixels by thresholding input observed images. This saturation correction leads to better reconstructions but the author does not deal with threshold value optimization. Yuan an et al. [15] tries to handle outliers by using a coarse-to-fine progressive deconvolution approach. This method can suppress ringing artifacts but still have disadvantages in output image.

Moreover, the kernel size in deconvolution also has significant influence in image quality and there is no guidance on how to find the optimal size for kernel in previous works. Joshi et al. [11] use inertial measurement sensors to estimate the motion of the camera over the course of the exposure, the experiment results show a clear improvement over the input blurry image but there is still some residually ringing that is unavoidable due to frequency loss during blurring. Stefan Harmeling et al. [17] introduced taxonomy of camera shake and constructed a method for blind deconvolution in the case of space-variant blur based on a recently introduced framework for space-variant filtering. The

disadvantage of method is that it can fail if the blurs are too large or if they vary too quickly across the image. Michael Hirsch et al. [16] combine Projective Motion Path Blur models and Efficient Filter Flow framework to develop a forward model for non-uniform camera shake. However, that method does not work effectively with moving or deformable objects.

To solve above-mentioned problems, in this paper, a new method for restoring the latent image from camera shake blur image is proposed by exploiting recent research in natural image statistics and obtained the improvement in experimental results. An algorithm is developed that explicitly handles outlier combine with an average difference criterion for estimating the blur kernel in the deconvolution process. A multi-scale approach is used in kernel estimation and some useful information in outlier modeling is exploited for reconstruct effectively the latent image. Our method is implemented with assuming the scene to be static, i.e., only the camera moves, and none of the photographed objects (no object motion). The problem of “blind” deblurring, where only a single blurry image is available is considered with assuming that blur is uniform blur and camera rotation is negligible. The model within the framework proposed by R. Fergus [7] for the PSF estimation case and S. Cho [[3]] for the case of non-blind deblurring is applied in this method.

II. MATERIALS AND METHODOLOGY

Blind deconvolution is typically addressed by first estimating the kernel, and then estimating the sharp image when the kernel is known. This method is not an exception. The deblurring implementation is classified into two main steps. In the first step, the blur kernel is estimated in a coarse-to-fine fashion in order to avoid local minima. The non-blind deconvolution with the outlier handling algorithm [3] for estimating the latent image is applied in the second step.

2.1. Image and Blur Model

The blur model can represent a blurred image b as a convolution of the latent image l with a blur kernel (as a Point Spread Function - PSF) k plus additional image noise n :

$$b = k \otimes l + n \quad (1)$$

where \otimes is the convolution operator, n denotes sensor noise at each pixel. The problem of blind-deconvolution is to recover the latent image l from blurred image b without specific knowledge of blur kernel k .

Recent research have shown that, the image statistics obey heavy-tailed distributions in their gradients, i.e. the gradients' distribution has most of its mass on small values but gives significantly more probability to large values than a Gaussian distribution. For estimating the latent image without the blur knowledge, the distribution over gradient magnitudes is represented with a zero-mean mixture-of-Gaussians model. This presentation can provide a good approximation to the empirical distribution, while allowing a tractable estimation procedure for our algorithm.

2.2. Blur Kernel Estimation

According to (1), given blurred image b , the posterior distribution is presented with Bayes' rule:

$$p(k, \nabla l | \nabla b) \propto p(\nabla b | k, \nabla l) p(k) p(\nabla l) \quad (2)$$

where ∇ denotes the gradient operator, k is blur kernel, ∇l is latent image gradient and ∇b is blur image gradient.

Our purpose is to approximate the full posterior distribution $p(k, \nabla l | \nabla b)$ and then estimate the kernel k with maximum marginal probability. The kernel is selected that is most likely with respect to the distribution of possible latent images, thus avoiding the overfitting that can occur when selecting a single “best” estimate of the image.

The prior of latent image gradients $p(\nabla l)$ is a mixture of zero-mean Gaussians, which is obtained based on Gaussian prior distributions:

$$p(\nabla l) = \prod_i \sum_{c=1}^C \pi_c N(\nabla l_i | 0, \nu_c) \quad (3)$$

where i indexes over image pixels, C denotes zero-mean Gaussians model, π_c and ν_c is weight and variance for the c -th Gaussian respectively, N is Gaussian distribution. $C = 4$ were used in our experiment.

The zero values in the kernel are encouraged by using sparsity prior $p(k)$:

$$p(k) = \prod_j \sum_{d=1}^D \pi_d E(k_j | \lambda_d) \quad (4)$$

where j indexes over blur kernel elements, D denotes exponential distribution with scale factors λ_d and weights π_d for the d -th component, E denotes exponential distribution. $D = 4$ were used in our experiment

Combination of (3) and (4) with assuming the plus noise is Gaussian can obtain:

$$p(\nabla b | k, \nabla l) = \prod_i N(\nabla b_i | k \otimes \nabla l_i, \sigma^2) \quad (5)$$

where i indexes over image pixels, N is Gaussian distribution, σ^2 is unknown noise variance.

For solving Eq.(2) using variational Bayesian approach, use Kullback-Leibler divergence to measure distance between the approximating distribution $q(k | \nabla l)$ and the true posterior $p(k, \nabla l | \nabla b)$, and the cost function C_{KL} is defined as:

$$C_{KL} = KL\{q(k, \nabla l, \sigma^2) \| p(k, \nabla l | \nabla b)\} - \ln p(\nabla b) \\ = \int q(\nabla l) \ln \frac{q(\nabla l)}{p(\nabla l)} d\nabla l + \int q(k) \ln \frac{q(k)}{p(k)} dk + \int q(-\sigma^2) \ln \frac{q(-\sigma^2)}{p(-\sigma^2)} d(-\sigma^2) \quad (6)$$

Because σ^2 is unknown, $q(k, \nabla l, \sigma^2)$ is used to represent $q(k, \nabla l)$. Minimize Eq.(6) by using iterative method based on variational Bayesian expectation maximization rule, the blur kernel can be estimated. The blur kernel k is estimated by varying image resolution in a coast-to-fine way, which starts with 3×3 kernel at coarsest level. The initial estimate for the latent gradient image is then produced by running the inference scheme, while holding k fixed. The pyramid running the inference at each level is then backup; the converged values of k and ∇l being up-sampled to act as an initialization for inference at the next scale up. At the finest scale, the inference converges to the full resolution kernel k .

2.3. Non-blind Deconvolution

After obtaining estimated and fixed the kernel, the final shape image is recovered by using non-blind deconvolution method proposed by Sunghyun Cho et al. [3].

Previous non-blind deconvolution methods assume a linear blur model where the blurred image is generated by a linear convolution of the latent image and the blur kernel. This assumption often does not hold in practice due to various types of outliers in the imaging process. Without proper outlier handling, recent methods may generate results with severe ringing artifacts even when the kernel is estimated accurately.

In previous work, Cho, S. et al. [3] demonstrate that outliers violate the linear blur assumption and consequently cause severe ringing artifacts to the result image. It is inappropriate to use a linear blur model when outliers exist, so it is necessary to avoid the violation of outliers, but most sources of the outliers are inevitable unfortunately. The method that masks out the outliers, Harmeling et al. [9], involves a threshold which distinguishes the outliers for masking out, but there is no guidance on how to find the optimal threshold value, so that method is not robust enough. Yuan et al. [15] proposed a directly suppressing artifacts approach, that method actually handle the outliers implicitly. An EM method which handles the outliers directly is proposed, and that is more efficient.

In this study, the MAP model is used for computing the most probable latent image l :

$$L = \arg \max p(l | k, b) \quad (7)$$

According to the Bayesian theory:

$$L = \arg \max_l \sum_{r \in R} p(b | r, k, l) p(r | k, l) p(l) \quad (8)$$

where R is the space of all possible configurations of r .

The latent image prior $p(l)$ is formulated as:

$$p(l) = \exp(-\lambda\phi(l))/Z \quad (9)$$

where Z is a normalization constant. Using the sparse prior, $\phi(l) = \sum_i \{ |(\nabla^h l)_i|^\alpha + |(\nabla^v l)_i|^\alpha \}$ were

used. In this case, $\alpha = 0.8$

Using EM method for solving equation (8) can obtain:

$$L_{E-\log} = E[\log p(b|r, k, l) + \log p(r|k, l)] \quad (10)$$

With assuming the noise is spatially independent, the likelihood $p(b|r, k, l) = \prod_i p(b_i|r, k, l)$, where:

$$p(b_i|r, k, l) = \begin{cases} N(b_i|f_i, \sigma) & \text{if } r_i = 1 \\ C & \text{if } r_i = 0 \end{cases} \quad (11)$$

where $f = k * l$, N is a Gaussian distribution, σ is the standard deviation. C is a constant defined as the inverse of the width of the dynamic range in the input image.

According to the prior model, with assuming r is spatially independent, $p(r|k, l) = \prod_i p(r_i|f_i)$, the

$p(r_i = 1|f_i)$ is defined:

$$p(r_i = 1|f_i) = \begin{cases} P & \text{if } f_i \in R \\ 0 & \text{if } f_i \notin R \end{cases} \quad (12)$$

where R is dynamic range, $P \in [0, 1]$ is the probability that pixel i is inlier. $R = [0, 1]$ were used for our implementation.

Putting (11), (12) into (10) can obtain:

$$L_{E-\log} = -\sum_i \frac{E[r_i]}{2\sigma^2} |b_i - f_i|^2 \quad (13)$$

where $E[r_i] = p(r_i = 1|b, k, l^0)$, By the Bayes' theorem, combination of (11) and (12) can obtain:

$$E[r_i] = \begin{cases} \frac{N(b_i|f_i^0, \sigma)P}{N(b_i|f_i^0, \sigma)P + C(1-P)} & \text{if } f_i^0 \in R \\ 0 & \text{if } f_i^0 \notin R \end{cases} \quad (14)$$

$E[r_i]$ is approximate 1 if the observed pixel i is inlier, and otherwise $E[r_i] = 0$

Use M step for finding the revised estimate L such that:

$$L_{\text{output}} = \arg \max_l \{ L_{E-\log} + \log p(l) \} \quad (15)$$

The values which computed in E step ($E[r_i]$) are used as pixel weights in the deconvolution process.

That process is performed in the M step. As a result, only inliers with large weights are used for deconvolution in the M step, while outliers with low weights are excluded.

The iteratively reweighted least squares (IRLS) method is used for solving (15), which is equivalent to minimizing:

$$L = \sum_i \omega_i^r |b_i - (k * l)_i|^2 + \lambda \left\{ \sum_i \{ \omega_i^h |(\nabla^h l)_i|^2 + \omega_i^v |(\nabla^v l)_i|^2 \} \right\} \quad (16)$$

where $\omega_i^r = E[r_i]/2\sigma^2$, $\omega_i^h = |(\nabla^h l)_i|^{\alpha-2}$, $\omega_i^v = |(\nabla^v l)_i|^{\alpha-2}$, finally, the latent gradient image can be get by alternating between updating $(\omega_i^h; \omega_i^v)$ and minimizing (16).

2.4. Kernel size optimization

In proposed method for PSF estimation and deconvolution, the PSF and deblurred image are obtained from input blur image but we don't know which size of PSF is better, and when output deblurred image

will reach the desired quality. This is also an ill-posed problem in recent research about blind deconvolution. The quality of output image depends on size of PSF. The optimal size of PSF (corresponding to the best quality of image) is different in each blur image. Hence, we propose a global multi-scale method for optimize the kernel by estimating the average difference parameters. This parameter is used because it is simple to calculate, have a clear physical meaning, and is mathematically convenient in the context of optimization.

The Average Difference parameter (AD) is used for estimating output image:

$$AD = \left| \sum_{j=1}^M \sum_{k=1}^N (x_{j,k} - x'_{j,k}) / MN \right| \quad (17)$$

Where M x N is size of blur image and deblurred image, after deconvolution operation.

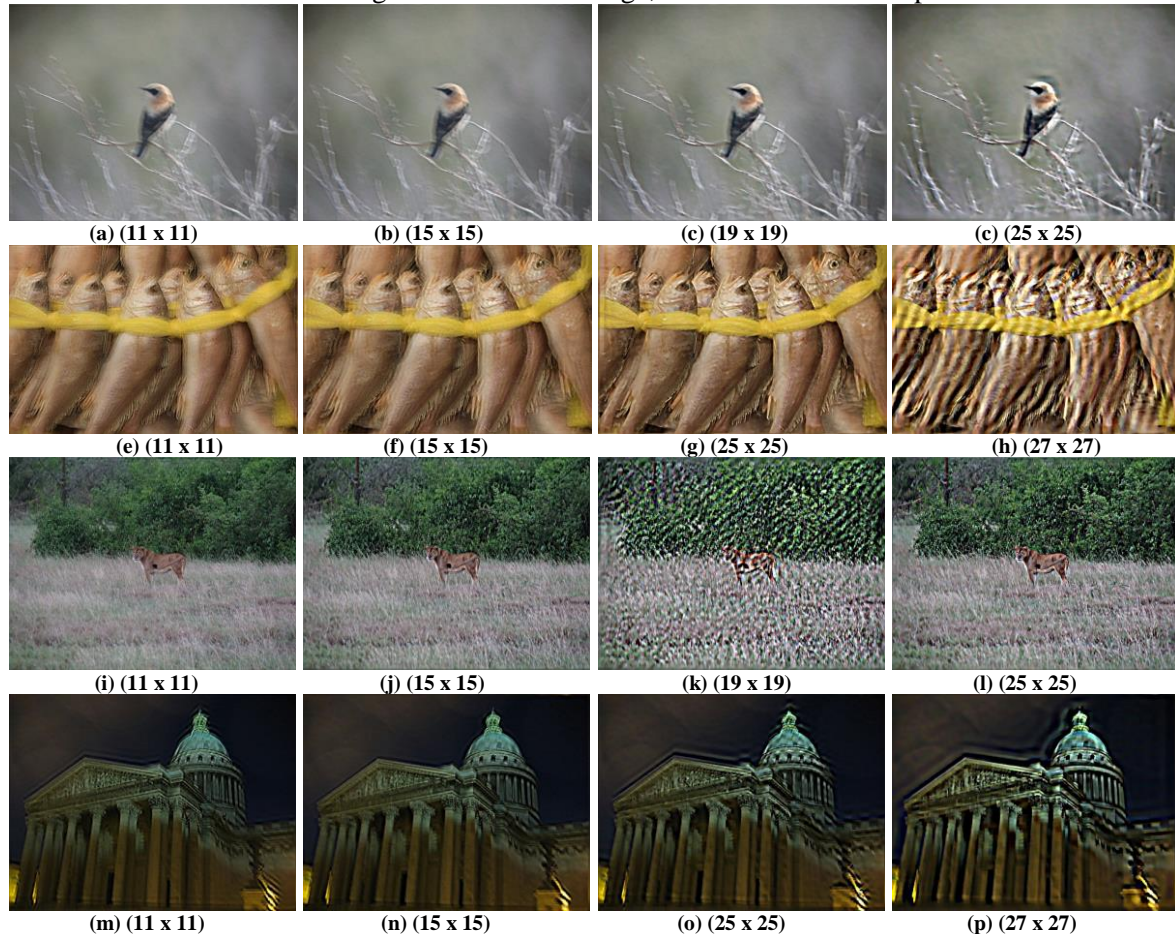


Figure 1. Experimental Results with different size of kernels

The image quality is estimated by evaluating the average difference value in comparison of output image and input blurred image. In practice, we found the output deblurred image can reach highest quality with minimum value of average difference (AD), and the ringing artifacts in deblurred image are more visible if the average difference (AD) is at a large value, as illustrated in Tab. 1 and Fig. 1. In practice, the output deblurred image can reach highest quality with minimum value of Average Difference, and the ringing artifacts in deblurred image are serious at large value of AD, as illustrated in Tab. 1 and Fig. 1.

The results of deconvolution with different PSFs are present in Fig. 1. Figs. 1(c), (g) (i), (n) are output deconvolution with (19x19) kernel in “Bird.jpg”, (25x25) kernel in “Fishes.jpg”, (11x11) kernel in “Lion.jpg” and (15x15) kernel in “Palace.jpg” respectively. Clearly these images are deblurred better than others that are deblurred with another size of kernel, the ringing artifacts are suppressed in these cases.

Table 1. The value of Average Difference with different size of kernels

PSF Size	Bird.jpg	Fishes.jpg	Lion.jpg	Building.jpg
11*11	0.0200	0.0293	0.0668	0.4275
13*13	0.0214	0.0394	0.0786	0.3364
15*15	0.0126	0.0501	0.0128	0.1929
17*17	0.0185	0.0532	0.5243	0.4275
19*19	0.0043	0.0552	4.2777	0.3108
21*21	0.0047	0.0247	2.7862	0.4861
23*23	0.0036	0.0263	0.3460	0.7682
25*25	0.0157	0.0080	0.6935	1.0469
27*27	0.2444	0.0162	1.0194	1.2087
29*29	0.2894	0.1040	1.2385	1.5126
31*31	0.3148	0.1040	2.3672	1.8504

III. EXPERIMENTAL RESULTS AND COMPARISONS

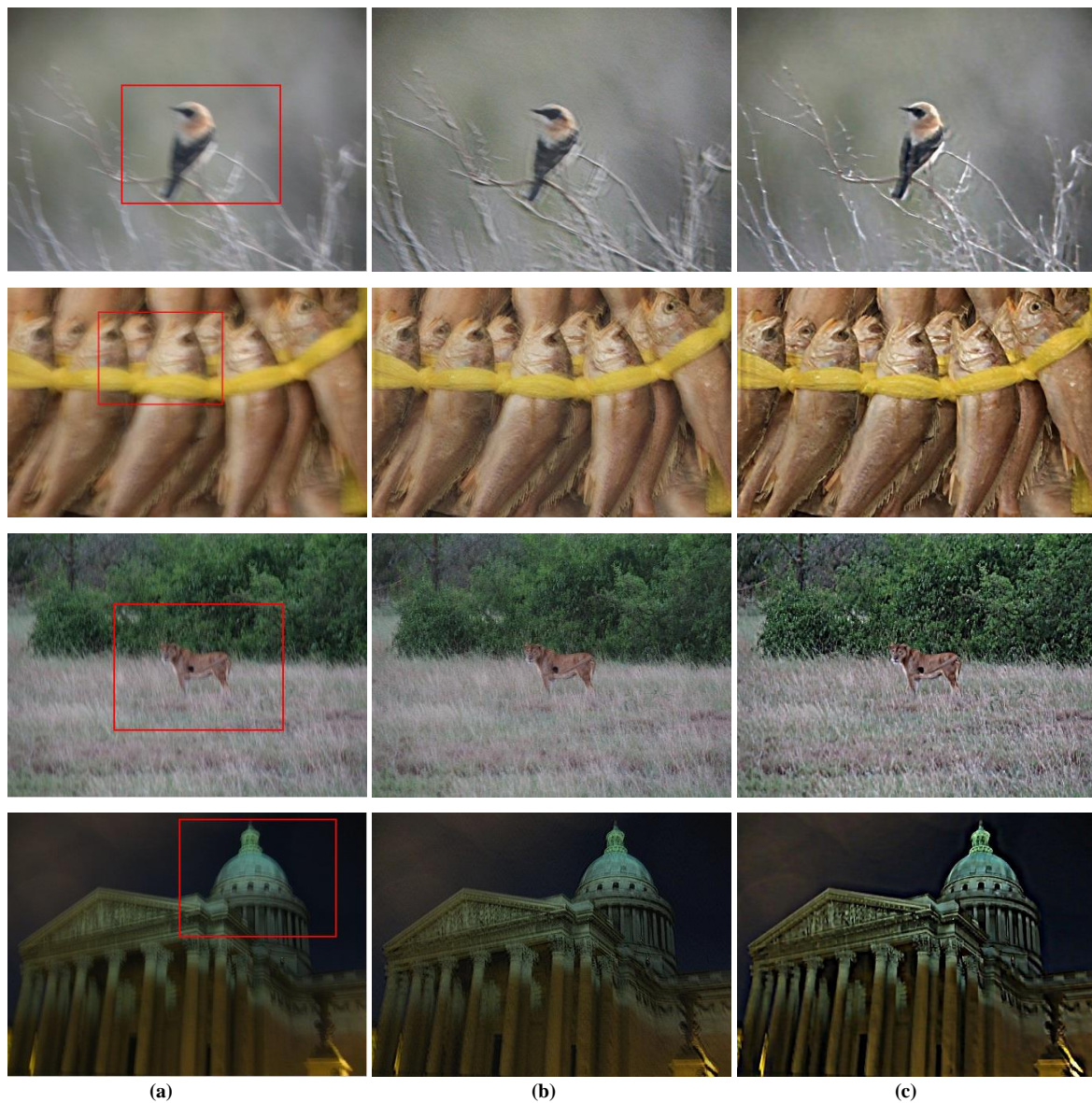


Figure 2. Comparisons to the method of Fergus (b)

The proposed algorithm and R. Fergus's algorithm [7] are demonstrated with the same size of blur kernel and have given some other results for comparisons. The input observed images are shown in

Figs. 2(a) (first column). The restored images by algorithm in [[7]] are displayed in Figs. 2(b) (second column), and the images restored by our algorithm are shown in Figs. 2(c) (third column) respectively. The image deconvolution is estimated by (19×19) , (25×25) , (15×15) and (15×15) size of kernel for "Bird", "Fishes", "Lion" and "Building" image, respectively. According to the restoration results, the proposed algorithm can recover the image quite well. Visually, we find that the restored images (third column) are clearer, brighter than those given in R. Fergus's algorithm [7].

The shadow around the bird is reduced in Fig. 2(b) (R. Fergus's Algorithm) and those effects have almost disappeared in Fig. 2(c) (our algorithm). The tree branches in Fig. 2(c) are also clearer than those branches in Figs. 2(a) and (b). The shaken effect is also suppressed completely in other pictures in Fig. 2(c) (third column). In addition, edges of objects are shown clearly in our results.

The improvement of method is exposed in close-up patches of the restored images, which are shown in Fig. 3. The close-up patches of restore images by method in [7] and our method are displayed in Figs. 3(a), Figs. 3(b) respectively. The edges are preserved well in Fergus's method but there has been the severe ringing and noise in output image. When the ringing and noise are reduced significantly (by increasing kernel size), the image details are also reduced. In contrast, our method show the accurate results with reduced ringing and noise in close-up patches, while preserving image edges well. The objects in image as shown in Figs. 3(b) are clear and the artifacts such as ringing and noise in the background are reduced significantly with our method. Again it is clear that the proposed algorithm can restore images quite well.

The benefit of handling outlier is illustrated in Fig. 4, where the pepper and salt noise is mixed in a blurry image that simulate the outliers in blur images, such as saturated pixel, red noise, dead pixel. In another image which was taken at night, the longer exposure time, the strong light source and the clipping function lead to the outliers in the image. In the analysis of Fig. 4(b), the salt and pepper noise make the kernel estimation fail, the Fergus's algorithm fails to obtain a proper kernel, that lead to bad results. In the analysis of Fig. 4(c), there exist serious ringing effects in the result image, the noise still exists and the restoration effect is not good enough. Actually, in the algorithm of paper [[18]], there exist pre-smooth process, that is Gaussian filter which reduce the pepper and salt noise and benefit to the restoration effect, but in the real photo-taking, some outlier (like Saturated pixel) have useful information. Filtering this kind of outlier simply is not a good method and can also lead to loss of information. In the analysis of Fig. 4(d), the experimental results are clearer, with the ringing effect reduced significantly, the salt and pepper noise faded away. In our algorithm, part of outlier information is used for rebuilding of that pixel and its surrounding pixels. Analyze the Fig. 4(f),(g),(h) in the same way, it is obvious that serious ringing effects and distortion exists in Fig. 4(f). There exists ringing effects and distortion on the top of the light area in Fig. 4(g), but in the results of our algorithm, the latent image is clear, the light area is rebuilt ideally, and there is no ringing effects and distortion.

The closed-up patches of the restored images are shown in Fig. 3. The closed-up patches of restore images by method in [7] and our method are displayed in Figs. 3(a), (c), (e), (g), Figs. 3(b), (d), (f), (h) respectively. Again it is clear that the proposed algorithm can restore images quite well.

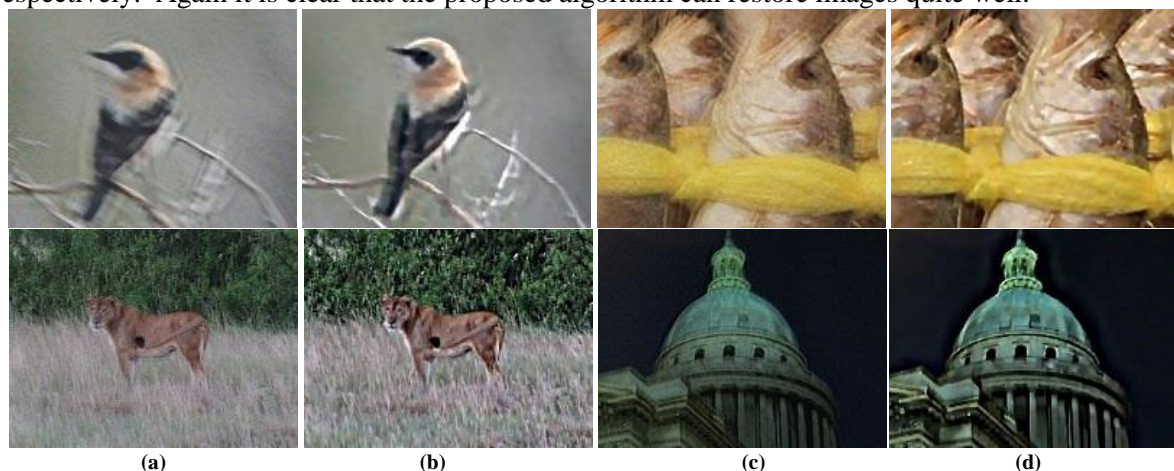


Figure 3. Closed-up patches of the restored images and comparisons to the method of Fergus (a, c)

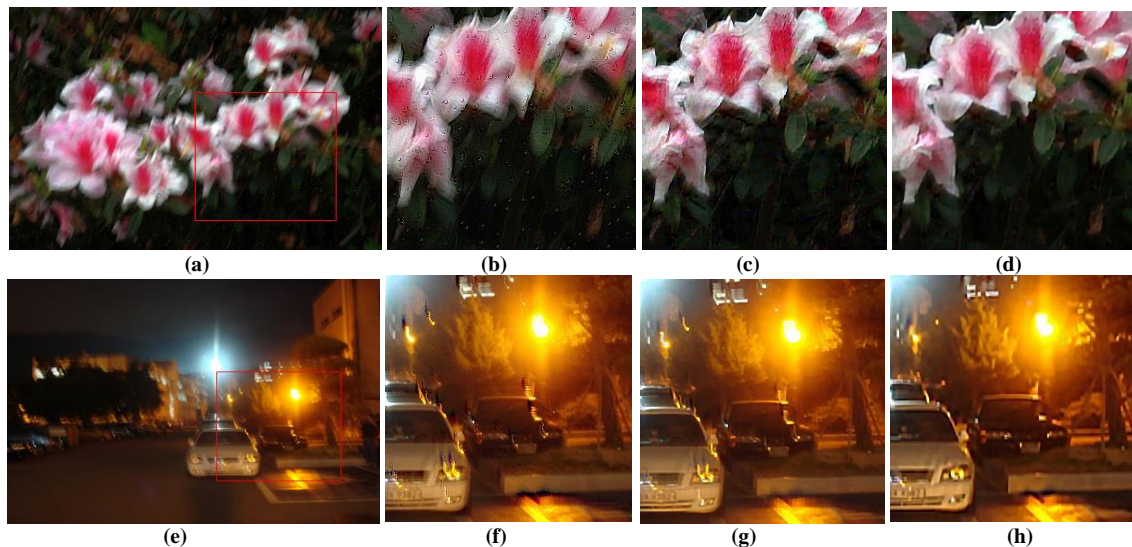


Figure 1. Closed-up patches of the restored images and comparisons with outlier handling

IV. CONCLUSION

In this study, an accuracy enhancement method for removing camera shake effects from photographs has been proposed. The proposed method can not only increase the robustness of kernel estimation but also obtain an accuracy quality of latent image. However, by applying natural image priors and advanced statistical techniques, plausible results can nonetheless be obtained. Such an approach may prove useful in other computational photography problems.

The results of our method often contain artifacts; most prominently, ringing artifacts occur near saturated regions and regions of significant object motion. We suspect that these artifacts can be blamed primarily on the non-blind deconvolution step. We believe that there is significant room for improvement by applying modern statistical methods to the non-blind deconvolution problem in future works.

REFERENCES

- [1] Caron, J., Namazi, N., Rollins, C. (2002). Noniterative blind data restoration by use of an extracted filter function, *Applied Optics* 41, 32 (November), pp 68–84.
- [2] Chen, J., Yuan, L., Tang, C. K. and Quan L. (2008). Robust dual motion deblurring, In *Proc. CVPR*.
- [3] Cho, S., Wang J., Lee S. (2011), Handling Outliers in Non-Blind Image Deconvolution, *IEEE Conf.*, pp. 495–502.
- [4] Cho, S., Matsushita, Y., Lee, S. (2007). Removing non-uniform motion blur from images, In *Proc. ICCV* 2007, pp. 1–8.
- [5] Cho, S. and Lee, S. (2009). Fast motion deblurring. *ACM Trans. Graphics (Proc. SIGGRAPH Asia* 2009), 28(5):145:1–145:8.
- [6] Dabov, K., Foi, A., Katkovnik, V., and Egiazarian, K. (2008). Image restoration by sparse 3D transform-domain collaborative filtering. *SPIE Electronic Imaging*.
- [7] Fergus, R., Singh, B., Hertzmann, A., Roweis, S.T., Freeman, W. (2006). Removing camera shake from a single photograph, *ACM T. Graphic.*, 25, pp. 787–794.
- [8] Gupta, A., Joshi, N., Zitnick, C. L., Cohen, M. and Curless, B. (2010). Single image deblurring using motion density functions, In *Proc. ECCV*.
- [9] Harmeling, S., Sra, S., Hirsch, M., Scholkopf, B. (2010). Multiframe blind deconvolution, super-resolution, and saturation correction via incremental EM. In *Proc. ICIP* 2010, pp. 3313–3316.
- [10] Joshi, N. (2008). Enhancing photographs using content-specific image priors. PhD thesis, University of California, San Diego.
- [11] Joshi, N., Kang, S. B., Zitnick, C. L. and Szeliski R. (2010). Image deblurring using inertial measurement sensors. *ACM Trans. Graphics (Proc. SIGGRAPH* 2010), 29(4):30:1–30:9.
- [12] Levin, A., Weiss, Y., Durand, F., Freeman, W. T. (2009). Understanding and evaluating blind deconvolution algorithms. *CVPR*.

- [13] Miskin, J., Mackay, D. J. C. (2000). Ensemble Learning for Blind Image Separation and Deconvolution, Adv. in Independent Component Analysis, M. Girolani, Ed. Springer-Verlag.
- [14] Shan, Q. , Jia, J., Agarwala, A. (2008). High-quality motion deblurring from a single image, SIG-GRAPH.
- [15] Yuan, L., Sun, J., Quan, L., Shum, H.Y. (2008). Progressive inter-scale and intra-scale non-blind image deconvolution, ACM Trans. Graphics.
- [16] Michael Hirsch, Christian J. Schuler, Stefan Harmeling, Bernhard Scholkopf (2011). Fast Removal of Non-uniform Camera Shake, Proc. IEEE International Conference on Computer Vision 2011.
- [17] Stefan Harmeling, Michael Hirsch, Bernhard Scholkopf (2010). Space-Variant Single-Image Blind Deconvolution for Removing Camera Shake, NIPS Conference 2010.
- [18] Li Xu, Jiaya Jia (2010). Two-Phase Kernel Estimation for Robust Motion Deblurring, ECCV 1, volume 6311 of Lecture Notes in Computer Science, page 157-170. Springer

AUTHORS

Quang Thi Nguyen

Male, was born in 1980. He received the B.S., M.S. degrees from the Le Qui Don Technical University, Hanoi, Vietnam, in 2004 and 2008, respectively. Now he is pursuing Ph.D. degree in School of Electronics and Information Engineering, Changchun University of Science and Technology. He currently focuses on Blind Deconvolution, Image processing.



Jun Xi Sun

Male, was born in 1971. He received the B.S., M.S. degrees from the Changchun University of Science and Technology, in 1995 and 2000, respectively, and the Ph.D Degree in Shanghai Jiao Tong University, in 2004. He is currently a Professor of School of Electronics and Information Engineering, Changchun University of Science and Technology. He currently focuses on Electronic Imaging, Digital Signal Processing.

Yang Sun

Male, was born in 1988. He received the B.S. degree from the Changchun University of Science and Technology, in 2011. Now he is pursuing M.S. degree in School of Electronics and Information Engineering, Changchun University of Science and Technology. He currently focuses on Blind Deconvolution, Image processing.