MOTION ESTIMATION AND DEBLURRING OF FAST MOVING OBJECTS

Jan Kotera, Filip Šroubek

Czech Academy of Sciences, Institute of Information Theory and Automation, Prague, Czech Republic

ABSTRACT

Image deblurring is one of the standard problems in image processing. Recently, this area of research is dominated by blind deblurring, where neither the sharp image nor the blur are known. The majority of works, however, target scenarios where the captured scene is static and the blur is caused by camera motion, i.e. the whole image is blurred. In this work we address a similar yet different scenario: an object moves in front of a static background. Such object is blurred due to motion while the background is sharp and partially occluded by the object. The problem of blind deblurring in such setting has not been properly addressed in literature. We formally define the problem, discuss its solvability, and explain why it cannot be viewed as a special case of classical blind deblurring. We propose a solution to the presented problem for a particular class of motions and demonstrate results on real data.

Index Terms— blind deblurring, object deblurring, motion estimation, alternating minimization

1. INTRODUCTION

Blind image deblurring (BD), the problem of removing blur from the observed image without knowing the blur itself, received considerable attention by the image processing and computer vision community in the last two decades. Methods based on various principles have been proposed and several surveys compiled [1, 2]. Early approaches considered space-invariant cases in which the same blur degrades the whole image [3, 4, 5, 6, 7, 8]. Later, space-variant scenarios were tackled, such as blur induced by complex camera motion [9, 10, 11]. Less frequently studied scenario is the multichannel BD [12, 13, 14], which assumes multiple differently blurred images of the same scene. Still, majority of works target scenario where static scene is blurred due to camera motion or incorrect focus and very few attempts include blur due to relative motion of objects and background. This is a surprising omission considering that objects in motion commonly appear in diverse real-world images or video sequences.

The notion of a fast moving object (FMO) was introduced in the context of object tracking in [15] as an object that moves over a distance exceeding its size within the exposure time. In this work we are interested in image deblurring, we therefore relax the definition and denote by FMO any bounded object that non-negligibly moves over static background. We refer to the task of restoring the FMO appearance and motion from the input image as the *FMO deblurring problem*. Resemblance to BD is evident, yet the two problems differ in several important ways. In particular, let us emphasize that FMO deblurring cannot be viewed as a special case of space-variant deblurring. The acquisition model implies that the FMO image and background are blended together and the background is partially occluded by the FMO, which cannot be modeled by space-variant convolution. Also, in its full generality, the FMO problem deals with complex object motion (3D rotation), which is again outside the scope of space-variant BD. The FMO problem is undoubtedly an interesting spin-off of image deblurring with clear ties to real-world applications.

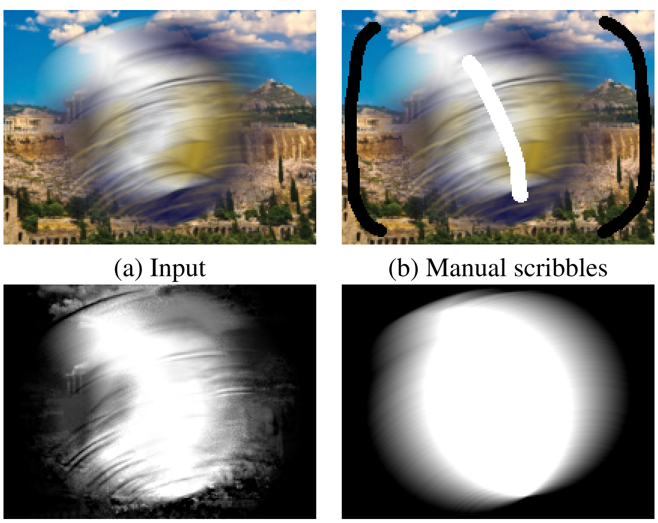
Image blur is commonly considered a nuisance, so the main goal of BD is to estimate the sharp image and the estimation of blur is only a necessary intermediate step. In the FMO problem, on the other hand, the blur could be of interest on its own since it contains information about the object motion. There are scenarios when the object appearance is known (e.g. in various ball sports) and estimating motion blur becomes of primary interest because it enables us to determine the object trajectory and rotation during exposure.

To our knowledge, the FMO deblurring problem has not been properly addressed in the literature, which makes this paper the first attempt in this direction. We focus on a simplification of the full FMO deblurring problem, namely a 2D object of known shape undergoing arbitrary motion and 2D rotation. Examples of such scenario is a flat object sliding on a table. Although the more general 3D case is principally the same, the increased computational and implementational complexity is significant and we do not include it here. Also, we assume that the background is known, which is not a restrictive assumption because the background image can be acquired separately or easily estimated from a video sequence.

Let us review some related work on motion deblurring similar to the FMO setting. One category of such methods considers the foreground-background blending due to motion and exploits the blurred object transparency map (alpha matte). Blind deconvolution of the transparency map is better

This work was supported by Czech Science Foundation grant GA18-05360S and SVV2017260452.

posed, since the latent sharp map is a binary image, which makes the blur estimation easier [16]. The same idea applied to rotating objects was proposed in [17]. An interesting variation was proposed in [18], where linear motion blur is estimated locally using a relation similar to optical flow. The main drawback of these methods is that they require a very accurate estimation of the transparency map as input, which is nearly impossible to achieve for typical FMO scenarios. The alpha matte estimation is traditionally aimed at segmentation-like tasks where most pixels are known to be background or foreground and very few pixels have fractional transparency. In addition, the color texture of both foreground and background needs to be simple [19], which is the key idea exploited in [16]. In the typical FMO input, however, the area of fractional transparency is large, the image textures complex and the alpha matte estimation ultimately fails despite manual help, see Fig. 1. The input image (a) contains ICIP logo moving over Athens background, in (b) scribbles are manually placed to indicate pure foreground and pure background, which is the required input for alpha matte estimation algorithms like [19] and difficult to do automatically. The result in (c) is still far too inaccurate for any subsequent PSF estimation, the true alpha matte is in (d).



(a) Input (b) Manual scribbles

(c) Output of [19] (d) True alpha matte

Fig. 1. Matte estimation results for FMO input. (a) Input image, (b) manually placed scribbles indicating foreground and background (requirement of the alpha matte estimation), (c) result of matte estimation by [19], (d) actual alpha matte.

Another approach is to ignore the blending model and instead estimate linear motion locally by exploiting the fact that autocorrelation increases in the direction of blur [20, 21]. Autocorrelation, however, requires large neighborhood to estimate blur parameters, which means that it is not suitable for small moving objects, non-uniform motion (rotation), and if the background blending is significant.

Interestingly, the FMO deblurring problem was mentioned in [22] as a rather marginal note, but the formulation is exactly the same as ours, though restricted to simple motions without rotations. The authors only theoretically suggested a semi-blind solution using hybrid camera setup and object tracking in a video sequence, it was not however carried out in practice. Applicability of such solution would be significantly limited to the proposed method.

The contribution of this work is as follows: We formulate the FMO acquisition model, corresponding deblurring problem, and present solution to the inverse problem to estimate the FMO appearance and motion from a single image and background. We discuss its properties and relationship to the BD problem. Finally we demonstrate the efficacy and some properties of the proposed method on several experiments on real data.

2. PROBLEM FORMULATION

Let the object with its 2D image denoted by f move in front of a static background b. Let the blurred appearance of f be the result of a linear blur operator H, that is Hf. For example, if the motion is pure translation then Hf will be convolution $Hf = h * f$ for some Point Spread Function h. As the object passes in front of the background, parts of the background get fully or partially occluded depending on how much time (as a fraction of exposure) any part of the object covers the particular background pixel. Let m denote the binary image of the object shape (indicator function of the object), then Hm is the fraction of exposure the object spends in front of each background pixel. As a result, we have the following model for the acquired image g, which describes the blending of f and b:

$$g = Hf + (1 - Hm) \cdot b. \tag{1}$$

The last product denoted by dot is pixelwise multiplication.

The blur operator H is characterized by a set of independent motion parameters h, as in the pure-translation case above where h is the PSF. More formally, if $(f, h) \to y$ is the bilinear blurring function of object f undergoing motion given by h, then H and F are the corresponding restrictions $H = (\cdot, h)$ and $F = (f, \cdot)$.

Assume that b and m are known, then the FMO deblurring problem consists of estimating the object appearance f and motion h from the input image g. To this end, we propose to solve the following minimization problem, inspired by work in classical deblurring

$$\min_{f,h} \frac{1}{2} \|Hf - (Hm) \cdot b - (g - b)\|_2^2 + \alpha_f \, \mathrm{TV}(f) + \alpha_h \|h\|_1$$
$$\text{s.t. } f_i \in [0, 1] \, \forall i \text{ and } h_i \geq 0 \, \forall i, \tag{2}$$

where $\mathrm{TV}(f) = \sum_i \|\nabla f_i\|_2$ is the total variation and i denotes the pixel index. We assume that the sought image f is in the $[0, 1]$ intensity range. As is common in BD, we solve this problem by alternatingly minimizing w.r.t. f and h. Below we highlight the most important aspects of the numerical solution.

2.1. Estimation of f

Minimizing (2) w.r.t. f amounts to solving

$$\min_f \frac{1}{2}\|Hf - b\cdot Hm - (g-b)\|_2^2 + \alpha_f \sum_i \|D_i f\|_2 + P_{[0,1]}(f), \tag{3}$$

where D_i denotes the gradient operator at the location i, and P is the set indicator function (infinite potential well) defined as

$$P_{[a,b]}(x) = \begin{cases} 0 & \text{if } x_i \in [a,b]\ \forall i, \\ +\infty & \text{otherwise.} \end{cases} \tag{4}$$

To solve this problem we use ADMM [23] with splitting $z_1 = Df$ due to the TV term and $z_2 = f$ due to the boundedness term with $u_{1,2}$ and ρ being the corresponding (scaled) duals and penalty weight, respectively. According to ADMM theory, we minimize (3) w.r.t. f and $z_{1,2}$ alternatingly and update $u_{1,2}$ in each iteration.

Minimization w.r.t. f is the linear system

$$\left(H^T H + \rho D^T D\right) f = H^T\left(b\cdot Hm + (g-b)\right) + \rho D^T(z_1 - u_1). \tag{5}$$

Minimization w.r.t. z_1 is soft-thresholding of $Df + u_1$, [23]. Minimization w.r.t. z_2 is projection of $f + u_2$ into $[0,1]$.

2.2. Estimation of h

Estimating h is similar to f estimation. The relevant subproblem of (2) is

$$\min_h \frac{1}{2}\|Fh - BMh - (g-b)\|_2^2 + \alpha_h \sum_i |h_i| + P_{[0,\infty]}(h), \tag{6}$$

where F and M are the linear operators performing blurring with fixed f, m resp. and B is the (diagonal) operator performing pixelwise multiplication with b. Again, we utilize ADMM with splitting $z = h$ due to ℓ^1 and positivity terms with u being the corresponding dual. The minimization w.r.t. h is the linear system

$$\left((F-BM)^T(F-BM) + \rho I\right) h = (F-BM)^T(g-b) + \rho(z-u). \tag{7}$$

Minimization w.r.t. z is again soft-thresholding of $h + u$ and truncation to $[0,\infty]$.

2.3. Rotational motion

The operator H can in principle represent blurring due to any kind of motion or shape/appearance change. In this paper, we limit ourselves to the case of translation and 2D rotation. Every motion blur can be approximated by a superposition of discretized orientations and positions

$$Hf = \sum_{ij} h_{ij} T_j R_i f = \sum_i H_i R_i f, \tag{8}$$

where T_j is a (fixed) translation operator into position j, R_i is a (fixed) rotation operator into orientation i and h_{ij} is a weight of the corresponding pose $T_j R_i$, i.e. h is the generalized blur that contains the unknown motion parameters. The second equality in (8) comes from the observation that the weighted sum of translations is convolution, $H_i = \sum_j h_{ij} T_j$, where H_i is a convolution operator. To summarize, we discretize the admissible rotation range into finite number of angles with corresponding R_i and then decompose H into these rotations each followed by standard convolution. The objective of motion estimation is to find h_{ij}, the convolution kernels h_i corresponding to each angle i.

Input Result Ground truth

Fig. 2. Deblurring result of real motion blurred object.

2.4. Remarks on problem solvability

For simplicity of the discussion, let us consider the pure-motion case, although the same arguments can be made with little modification in the rotational-motion case. Standard blind image deblurring solves the problem of finding f and h such that $h * f = g$, therefore minimizes the error term $\|h * f - g\|$. The problem is that there are solutions $(\hat{f}, \hat{h})$ that minimize the error term perfectly and yet are practically useless. The most striking example is the so-called no-blur solution $\hat{h} = \delta$, $\hat{f} = g$, where δ is the Dirac delta function. Then obviously $\hat{h} * \hat{f} = \delta * g = g$ and the error term vanishes (in the noiseless case), so this solution presents a strong (albeit possibly local) minimum and is one of the reasons why the functional minimization is practically problematic.

In the FMO deblurring problem, there are circumstances when this trivial solution, and as a consequence corresponding minimum of the functional in (2), do not exist. Let us assume that the object f has support (given by m) smaller than input image g, so that there is room for motion, and that the true motion h is non-trivial, $h \neq \delta$. Then the area affected by the object motion in the input image g is larger than the support of f. Let $\hat{h} = \delta$ be arbitrarily placed δ function, then there is no $\hat{f}$ such that the pair $(\hat{f}, \hat{h})$ satisfies eq. (1), because $\delta * \hat{f} - (\delta * m)\cdot b$ is nonzero only in the neighborhood of δ given by the support of f (equiv. m) while $(g - b)$ is nonzero in larger area due to the assumption of motion non-triviality.

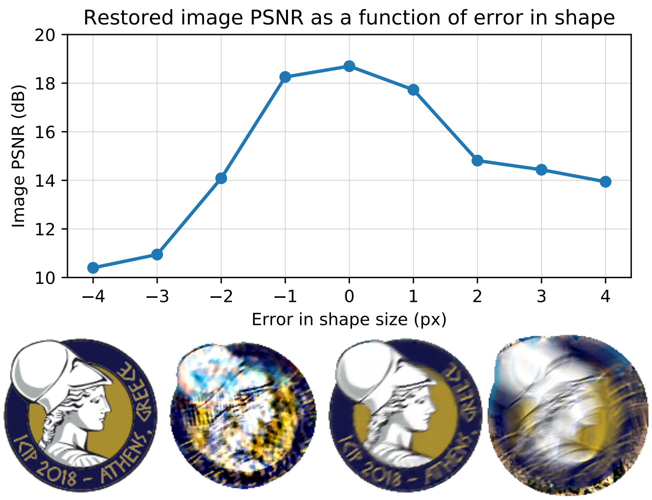
The discussion above is not exhaustive, but it nevertheless shows that in some aspects the FMO deblurring problem is actually easier to solve than classical BD.

3. EXPERIMENTAL VERIFICATION

We demonstrate the efficacy of our method on several real authentically blurred images. In the experiment we slid a picture

Input Result Ground truth

Fig. 3. Deblurring result of real motion blurred object.

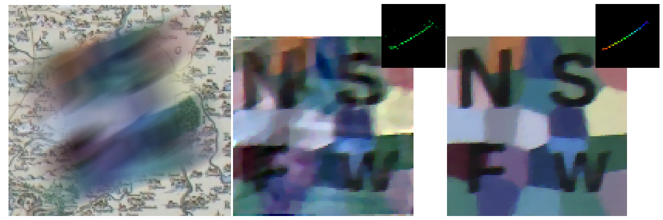


Orig. image $-$ 3px Correct size $+$ 3px

Fig. 4. Deblurring of the image in Fig. 1a with inaccurate object size. The upper plot shows the restored image PSNR when deblurred with smaller or larger prescribed object size than accurate. The bottom row shows the original object and some selected results.

on a table while capturing the table with a video camera. We then took one frame from the sequence as the input to deblurring and used the neighboring frames to determine the background image. The object shape m was estimated by simple background subtraction from one of the frames in which the object was at rest. The deblurring results are in Fig. 2-3. The simple image in Fig. 2 is restored quite well while the card in Fig. 3 lost some of its details. The estimated SNR of the camera is approx. 39dB, which is quite low, and it is one of the reasons why fine details are hard to recover after severe blurring.

The proposed method requires knowing the object shape m as an input to the deblurring. We are working on an extension to be published in the future which removes this requirement, it is nevertheless informative so see how critical the precise knowledge of the shape is for successful deblurring. For this purpose we prepared a synthetic experiment in which we deblur the image in Fig. 1a using slightly smaller or larger object mask and measured the quality of the result. The



(a) Input (b) Fixed rotation (c) With rotation estimation

Fig. 5. Deblurring with and without rotation handling. The input in (a) rotates by approx. 3° during exposure, the restored images PSNRs are 21dB and 28dB, resp. Superimposed are the estimated motion PSFs with color-coded orientation angle. Image in (b) is deblurred considering fixed mean orientation while in (c) the orientation is estimated as part of motion estimation.

results are in Fig. 4. The error of 1px in shape size (radius of the logo) constitutes approx. 1.5% of the object size, so we can see that the restoration quality deteoriates quickly as the size error increases. Some of the restored images as well as the original are in the bottom row of Fig. 4. We can conclude that accurate shape knowledge is quite important.

Equally important is correct motion handling, e.g. including rotation in the model even if it is very small. The image in Fig. 5a rotates by mere 3° during exposure and yet ignoring rotation in the deblurring results in visible loss in quality and approx. 7dB drop in PSNR, as seen in Fig. 5bc along with the estimated motion and rotation.

The runtime of the method is difficult to specify, as it depends strongly on the image and motion size and prescribed convergence criteria. As a reference, deblurring of 100px$\times$100px image with 30px long PSF and 15° rotation takes roughly 1 minute in our MATLAB implementation.

4. CONCLUDING REMARKS

We have presented a problem with practical applications that falls into the category of blind image deblurring and which has not been properly addressed by existing methods. We have also proposed a method to solve a simplified version of the problem, namely with known object shape and motion limited to 2D, and discussed some theoretical aspects of its solvability. Experiments on real data demonstrate the efficacy of the proposed method as well as some of its properties. The method is sensitive to noise and therefore works well particularly with images for which loss of detail due to noise is not very prominent. Accurate knowledge of the object shape as well as correct motion modeling prove to be critical, our future work therefore naturally leads in this direction. To make the proposed method more practically useful, we are working on an extension to estimate the object shape from the blurred input, include 3D motions and objects, model real-world effects such as illumination changes and camera perspective distortion, and increase robustness to noise.

5. REFERENCES

[1] P. Campisi and K. Egiazarian, Eds., *Blind Image Deconvolution, Theory and Application*, CRC Press, 2007.

[2] Pablo Ruiz, Xu Zhou, Javier Mateos, Rafael Molina, and Aggelos K. Katsaggelos, "Variational Bayesian blind image deconvolution: A review," *Digital Signal Processing*, vol. 47, pp. 116–127, dec 2015.

[3] Rob Fergus, Barun Singh, Aaron Hertzmann, Sam T. Roweis, and William T. Freeman, "Removing camera shake from a single photograph," in *SIGGRAPH '06: ACM SIGGRAPH 2006 Papers*, New York, NY, USA, 2006, pp. 787–794, ACM.

[4] R. Molina, J. Mateos, and A. K. Katsaggelos, "Blind deconvolution using a variational approach to parameter, image, and blur estimation," *IEEE Transactions on Image Processing*, vol. 15, no. 12, pp. 3715–3727, Dec. 2006.

[5] Li Xu and Jiaya Jia, "Two-phase kernel estimation for robust motion deblurring," in *Proceedings of the 11th European conference on Computer vision: Part I*, Berlin, Heidelberg, 2010, ECCV'10, pp. 157–170, Springer-Verlag.

[6] Li Xu, Shicheng Zheng, and Jiaya Jia, "Unnatural l0 Sparse representation for natural image deblurring," in *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, 2013, pp. 1107–1114.

[7] David Wipf and Haichao Zhang, "Revisiting bayesian blind deconvolution," *Journal of Machine Learning Research (JMLR)*, November 2014.

[8] Daniele Perrone and Paolo Favaro, "A logarithmic image prior for blind deconvolution," *International Journal of Computer Vision*, vol. 117, no. 2, pp. 159, Apr. 2016.

[9] M. Hirsch, C. J. Schuler, S. Harmeling, and B. Scholkopf, "Fast removal of non-uniform camera shake," in *Proc. IEEE Int Computer Vision (ICCV) Conf*, 2011, pp. 463–470.

[10] O. Whyte, J. Sivic, A. Zisserman, and J. Ponce, "Non-uniform deblurring for shaken images," *International Journal of Computer Vision*, vol. 98, no. 2, pp. 168–186, 2012.

[11] J. Pan, D. Sun, H. Pfister, and M. H. Yang, "Deblurring images via dark channel prior," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PP, no. 99, pp. 1, 2017.

[12] F. Šroubek and J. Flusser, "Multichannel blind iterative image restoration," *IEEE Trans. Image Processing*, vol. 12, no. 9, pp. 1094–1106, Sept. 2003.

[13] Alex Rav-Acha and Shmuel Peleg, "Two motion-blurred images are better than one," *Pattern Recogn. Lett.*, vol. 26, no. 3, pp. 311–317, 2005.

[14] Haichao Zhang, D. Wipf, and Yanning Zhang, "Multi-observation blind deconvolution with an adaptive sparse prior," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 8, pp. 1628–1643, 2014.

[15] D. Rozumnyi, J. Kotera, F. Šroubek, L. Novotný, and J. Matas, "The world of fast moving objects," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, July 2017, pp. 4838–4846.

[16] Jiaya Jia, "Single image motion deblurring using transparency," in *Computer Vision and Pattern Recognition (CVPR), 2007 IEEE Conference on*, 2007, pp. 1–8.

[17] Qi Shan, Wei Xiong, and Jiaya Jia, "Rotational motion deblurring of a rigid object from a single image," in *Proc. IEEE 11th International Conference on Computer Vision ICCV 2007*, Oct. 2007, pp. 1–8.

[18] Shengyang Dai and Ying Wu, "Motion from blur," in *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, June 2008, pp. 1 –8.

[19] A. Levin, D. Lischinski, and Y. Weiss, "A closed-form solution to natural image matting," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 2, pp. 228–242, Feb 2008.

[20] Tae Hyun Kim and Kyoung Mu Lee, "Segmentation-free dynamic scene deblurring," in *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*, 2014, pp. 2766–2773.

[21] Jian Sun, Wenfei Cao, Zongben Xu, and J. Ponce, "Learning a convolutional neural network for non-uniform motion blur removal," in *Computer Vision and Pattern Recognition (CVPR), 2015 IEEE Conference on*, 2015, pp. 769–777.

[22] S. K. Nayar and M. Ben-Ezra, "Motion-based motion deblurring," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 6, pp. 689–698, June 2004.

[23] Stephen Boyd, Neal Parikh, Eric Chu, Borja Peleato, and Jonathan Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, Jan. 2011.