IMPROVING ALPHA MATTING AND MOTION BLURRED FOREGROUND ESTIMATION

Rolf Köhler [1], *Michael Hirsch* [2,1], *Bernhard Schölkopf* [1], *Stefan Harmeling* [1]

[1]MPI for Intelligent Systems, Spemannstrasse 38, 72076 Tübingen, Germany
[2]University College London, Department of Physics and Astronomy, London, UK

ABSTRACT

We present a new method for separating motion blurred foreground objects from their background given a single image. Previous techniques focused on estimating alpha mattes for separating sharp, non-moving foreground objects from fairly homogeneous background. In those cases the only pixels which are ambiguous are those which exhibit fractional pixel occupancy. In this paper, we address the problem of alpha matte and foreground estimation of motion blurred objects. We show, that explicit modeling of the object motion facilitates the estimation and improves the quality of the estimated alpha mattes. In addition, we improve foreground extraction of motion blurred objects with a new regularization term. This task is particularly difficult in smeared out regions, where the background shimmers through. Both synthetic and real-world examples illustrate the merit of our approach.

***Index Terms*—** alpha matting, motion blur, foreground estimation.

1. INTRODUCTION

Motion blur is a common problem in photography as it causes image blur that destroys details in the captured photo. While camera shake (ego motion) is often a problem in hand-held photography, a moving object (object motion) during image capture leads to image blur even in case of a static camera. This typically happens when the object's speed is fast compared to the exposure time, such that different projections of the object hit the camera's sensor plane during the time the shutter is open. While ego motion affects the whole image, object motion leads to image blur only in those image regions that are affected by the object motion. Otherwise the photo is sharp. Segmenting an image corrupted by object motion into fore- and background is especially difficult in the smeared-out boundary region. There, the blurry foreground is partially transparent and the background shimmers through.

The common way to model a partially transparent boundary between foreground and background is via alpha matting, which models the observed image I as a combination of foreground $\tilde{F}$ and background $\tilde{B}$:

$$I = \alpha \odot \tilde{F} + (1 - \alpha) \odot \tilde{B} \tag{1}$$

Corresponding author Rolf Köhler can be reached by email at rolf.koehler@tuebingen.mpg.de.

where $0 \leq \alpha \leq 1$ is the alpha matte and $\odot$ the pixel-wise product.

Common alpha matting methods assume a static scene and do not model possible blur of the foreground due to object motion. However, in this paper we would like to take advantage of that additional knowledge, so we explicitly incorporate the blur into the forward model [1, 2, 3]:

$$I = k * (\beta \odot F) + (1 - k * \beta) \odot B. \tag{2}$$

k is the point spread function (PSF) of a spatially-invariant blur of the foreground and $*$ denotes the convolution operation. Note that the results easily generalize to spatially-variant blur.

How are α, $\tilde{F}$ and $\tilde{B}$ in (1) related to β, k, F and B in (2)? For this we first consider the terms of the background image:

$$(1 - k * \beta) \odot B = (1 - \alpha) \odot \tilde{B}. \tag{3}$$

which suggests to choose $\alpha = k * \beta$ and $B = \tilde{B}$. For the foreground expressions we get:

$$k * (\beta \odot F) = \alpha \odot \tilde{F} \tag{4}$$

which is fulfilled for $\tilde{F} = \alpha^{-1} \odot (k * (\beta \odot F))$ for all pixels where α is not zero. At the other pixel locations $\tilde{F}$ can be chosen arbitrarily since they are not relevant in Eq. (1). We see that an image modelled by Eq. (2) can be approached as a standard alpha matting problem (i.e. by Eq. (1)) if we constrain α to be $k * \beta$. In Sec. 3 we show that this constraint pushes the alpha matting algorithm of [4] towards the true solution.

2. RELATED WORK AND CONTRIBUTIONS

Most alpha matting techniques [4, 5, 6, 7, 8, 9] are designed to extract a sharp foreground, not a blurry one. For a benchmark data-set for this setup see [10].

There are several papers which consider alpha matting with motion-blurred foreground, often using additional hardware or making use of several images: [11] uses alpha matting as a preprocessing step to extract the blurry foreground from a low-frequency background, [12, 13, 14] use a hybrid camera for motion correction, [2] performs matting using multiple cameras with different focal settings, [15] uses several differently polarized images and [16] a camera array.

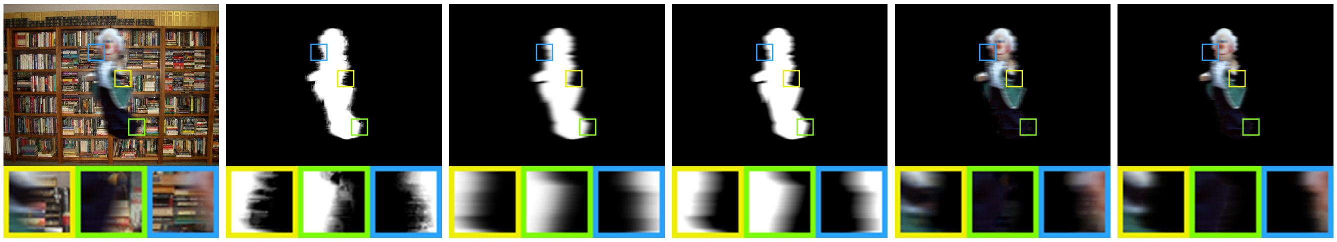


Fig. 1. Improving the alpha matting approach of [4] for images containing motion blurred foreground objects in the case of a synthetic example. The alpha matte of [4] shown in panel (b) suffers from strong artifacts at the boundaries, whereas our alpha matte (c) closely resembles the true matte (d). For fair comparison we used the ground truth matte (d) for the foreground estimation: the foreground estimated by [4] shows artifacts at the boundaries (e), whereas our foreground (f) is smooth in the boundary region. This figure is best viewed on screen.

All of the above mentioned approaches use several images or specialized hardware. There are some methods which try to improve alpha matting for motion blurred objects based on a single image. These use the motion blur indirectly to improve the alpha matting: [3] adds a regularization term to the energy function of [4] and [7]. This regularization term incorporates motion information into the matting cost function and promotes solutions which are smooth in the motion direction. [17] uses a guidance matte to improve the alpha matting results of [4]. They start with the alpha matte estimated by [4], which usually exhibits some artifacts. Those artifacts are removed by deblurring the alpha matte with the motion blur PSF, using a sparsity constraint and then blurring it again. This blurry matte is then incorporated as a soft constraint in the closed form matting [4]. In [18], the authors employ a smoothness and binary constraint to obtain an alpha matte which fits better to a motion blurred object than the jagged alpha matte obtained by [4], which the authors use as an initialization to their algorithm. [19] considers rotational motion blur and builds on user input to constrain the solution.

Recovering the foreground given the alpha matte is often not discussed in the alpha matting literature. For motion blurred foreground objects this is particularly difficult task, since there is usually a large transparent boundary region as can be seen e.g. in Fig. 5. [20] applies the unaltered approach of [4] which recovers the foreground by adding regularization terms which smooth the recovered foreground in the horizontal and vertical directions in regions with sharp boundaries in the alpha matte. [3] extends the regularization in [4] to eight directions, emphasizing the direction of the motion.

While the approaches discussed above incorporate the information about the motion blur in an *indirect* way, we modify the image generation model and adapt the algorithm of [4] to *directly* make use the motion blur model of Eq. (2). The advantage of this approach is that it explicitly models and exploits the motion blur information (Sec. 3). Furthermore, we improve the foreground recovery of [4] by giving hints about the correct color of the motion blurred object in the smeared out region (in Sec. 4).

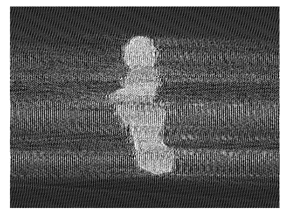


Fig. 2. Severe ringing artifacts arise when the alpha matte β is estimated without an L2 prior.

3. ESTIMATING THE MOTION BLURRED MATTE

The closed form solution by [4] is derived by assuming that both $\tilde{F}$ and $\tilde{B}$ are approximately constant over a small window around each pixel. The resulting cost function,

$$C(\alpha) = \alpha^T L\alpha + \lambda(\alpha - b_S)^T D_S(\alpha - b_S) \qquad (5)$$

is optimized wrt. the alpha matte α. It consists of two terms: the first is based on a Laplacian matrix L which expresses that close-by pixels with similar colors should be in the same region (i.e. either both in the background or both in the foreground). The second term includes the information from the trimap which defines known fore- and background regions given by user scribbles.

To demonstrate the short-comings of this approach for motion blurred foreground, we generated a toy example where we know the true k, β, B and F. Panel (b) of Fig. 1 shows that the unmodified approach in [4] leads to artifacts at the boundaries.

Let us now assume that we know, in addition to a user trimap, also the blur kernel k of the foreground, which could be obtained by blind deconvolution methods like [21, 22, 23] on a user-specified rectangle inside the foreground region. Instead of minimizing $C(\alpha)$ in α we minimize $C(k * \beta)$ in β. Note that this does not make the problem more difficult, since $k * \beta$ is linear in β, i.e. there exists a matrix K such that $K\beta = k * \beta$. Unfortunately, using this intuitive approach without any further modifications leads to ringing artifacts in β (see Fig. 2).

To reduce this ringing, we add a L_2 smoothness prior on

β, i.e. we minimize

$$C(k * \beta) + \gamma(\|D_x\beta\|^2 + \|D_y\beta\|^2) \quad (6)$$

in β with D_x and D_y being horizontal and vertical derivative operators (represented as matrices) and γ a regularization constant. Setting the derivative of Eq. (6) wrt. β to zero we obtain a linear system

$$[K^T(L+\lambda D_S)K+\gamma(D_x^T D_x+D_y^T D_y)]\beta = \lambda K^T D_S b_S \quad (7)$$

which can be solved using Matlab's backslash operator (similar to the implementation of [4]). Finally, the estimated β is blurred by the kernel k to obtain the alpha matte $\alpha = k * \beta$.

Fig. 1 shows the difference of an alpha matte using the matting algorithm by [4] and our approach. The matte by [4] is frayed on the boundaries whereas our matte is smooth as expected for a motion blurred object. Note, that in all experiments we set $\lambda = 100$ and $\gamma = 150$. We found that the exact choice of those parameters is not crucial for the algorithm to yield good results.

4. RECOVERING THE BLURRY FOREGROUND

Recovering the foreground, given the alpha matte is only an easy task if the background color is rather homogeneous and if the boundary regions with $0 < \alpha < 1$ are small. In images where the background is textured and contains high frequency content which partially shimmers through in the boundaries of the motion blurred region, the correct color of the foreground object is hard to estimate.

To compare different foreground recovery strategies, let's assume that the true alpha matte is given. [4] recovers foreground and background by minimizing the function

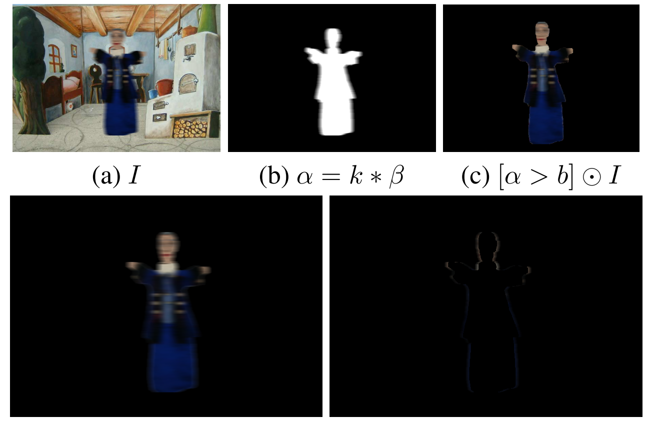
$$\sum_{i\in I}\sum_{c}(\alpha_i F_i^c + (1-\alpha_i)B_i^c - I_i^c)^2$$
$$+ |\alpha_{i_x}|\big((F_{i_x}^c)^2 + (B_{i_x}^c)^2\big) + |\alpha_{i_y}|\big((F_{i_y}^c)^2 + (B_{i_y}^c)^2\big), \quad (8)$$

where c denotes the different color channels. The first term follows from Eq. (1) and the second term is a constraint on the x and y derivatives $F_{i_x}^c, F_{i_y}^c, B_{i_x}^c$ and $B_{i_y}^c$ of F^c and B^c. α_{i_x} and α_{i_y} are the matte derivatives.

Applying this approach to the toy example using the ground truth matte α^*, results in several artifacts at the boundary (see Fig. 1(d)), which was also reported previously (e.g. Sec. 5 in [20]). To reduce these artifacts, we propose to add a soft constraint, which gives a hint on the color of the foreground in the smeared out boundary regions.

We observed that in the smeared out boundary regions the color of the foreground is often the same as the color of the blurry object (see Fig. 3). This can be expressed by a regularization term that we define with Iverson brackets $[\cdot]$, which abbreviate indicator functions, i.e. $[p] = 1$ if proposition p is true, and zero otherwise.

Given the matte $\alpha = k * \beta$, we first extract the regions of the foreground, i.e. where the alpha matte is larger than some threshold, $\alpha > b$. Choosing $b = 0.7$ leads to good results, which we use for all images shown in this paper. The expression $[\alpha > b] \odot I$ represents the foreground region that contains the colors which are blurred into the transparent boundary region of the foreground. By blurring this foreground region with the PSF k (resulting in $F_k = k*([\alpha > b]\odot I)$), we extend these colors into the transparent region, i.e. where $0 < \alpha < b$. Only pixels in that region are regularized. This is expressed by the following regularization term



(a) I (b) $\alpha = k * \beta$ (c) $[\alpha > b] \odot I$

(d) $F_k = k * ([\alpha > b] \odot I)$ (e) $[0 < \alpha < b] \odot F_k$

Fig. 3. Recovering the blurry foreground: (a) input image with blurry foreground, (b) the corresponding (blurry) alpha matte, (c) shows the part of the foreground where $\alpha > b$ with $b = 0.7$. Image (c) blurred with the true blur kernel k is shown in (d). Panel (e) shows the boundaries of image (d) for the region of alpha ($0 < \alpha < b$). It is evident that image (e) is a good soft constraint for recovering the correct foreground color in the difficult transparent region, where $0 < \alpha < b$.

$$R(\alpha) = \gamma\big\|[0 < \alpha < b] \odot (\tilde{F} - (k * ([\alpha > b] \odot I)))\big\|^2. \quad (9)$$

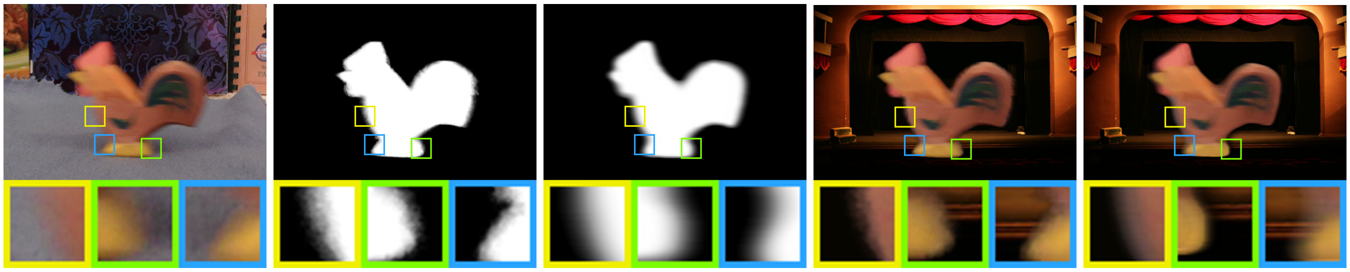
Combining this with Levin's approach (Eq. (8)) leads to improved recovery of the foreground region as can be seen in Fig. 1(e).

5. RESULTS

In the previous sections we have already shown in a number of synthetic experiments that our proposed modifications to [4] do improve the matting performance for motion-blurred foreground objects. In this section, we show results on real-world examples (taken by [3] and by ourselves).

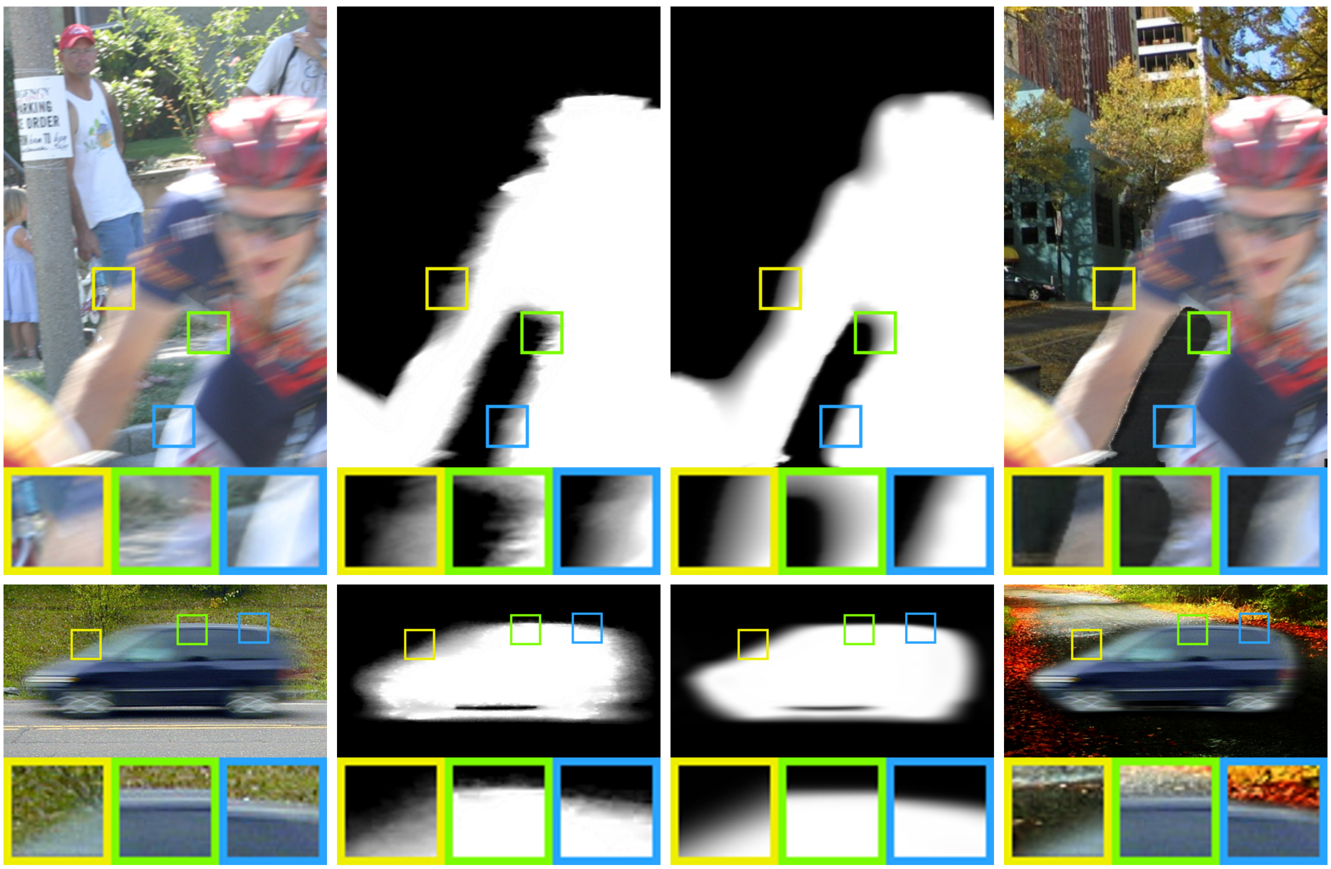
Fig. 4 compares our approach with [4] that does not use the PSF information. The alpha matte estimated by us is smoother and shows fewer artifacts (panels (b) vs. (c)). This also leads to a better boundary if we place the extracted foreground onto a new background (panels (d) vs. (e)).

Fig. 5 compares our approach with the method of [3], which is the only alpha matting method (as far as we know), that tries to improve alpha matting for motion blurred foreground objects. The close-ups (panels (b) vs. (c)) show that the estimated alpha matte of [3] suffers from artifacts,



(a) input (b) alpha by [4] (c) our alpha (d) foreground by [4] (e) our foreground

Fig. 4. Comparison with [4]: (a) input image with motion blurred foreground, (b) and (c) estimated alpha mattes, (d) and (e) composed example, where we pasted the estimated foreground of (a) onto a new background.



(a) images from [3] (b) alpha by [3] (c) our alpha (d) our foreground

Fig. 5. Comparison with [3]: (a) input images taken from [3], (b) and (c) show estimated alpha mattes, (d) composed example, where we pasted our recovered foreground onto a new background. Note that the recovered foreground of [3] was not available.

whereas our alpha matte is smooth as we would expect for an alpha matte of a moving object. We also manage to recover a smoother foreground. Especially in regions where the background shimmers through we obtain a reasonable color estimation.

6. CONCLUSION

We presented a new approach for alpha matting and foreground estimation of images containing motion blurred foreground objects. For this we extended the alpha matting algorithm of [4] by incorporating the motion blur in form of a point spread function k into the image generation model such that it can better estimate the alpha matte of a blurry foreground. As a second contribution we presented a novel regularization term which facilitates foreground estimation and leads to improved results, which we verified on both synthetic and real-world examples.

In future work, we will extend our approach to spatially varying motion blur (e.g. by using efficient filter flow framework of [24]).

7. REFERENCES

[1] A. Loktyushin and S. Harmeling, "Automatic foreground-background refocusing," in *IEEE International Conference on Image Processing (ICIP)*, 2011.

[2] M. McGuire, W. Matusik, H. Pfister, J.F. Hughes, and F. Durand, "Defocus video matting," in *ACM Transactions on Graphics (SIGGRAPH)*, 2005, pp. 567–576.

[3] H.T. Lin, Y.W. Tai, and M.S. Brown, "Motion regularization for matting motion blurred objects," *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, vol. 33, no. 11, pp. 2329–2336, 2011.

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

[5] A. Levin, A. Rav-Acha, and D. Lischinski, "Spectral matting," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2007, pp. 1–8.

[6] C. Rhemann, C. Rother, and M. Gelautz, "Improving color modeling for alpha matting," in *British Machine Vision Conference*, 2008, vol. 2, pp. 1155–1164.

[7] J. Wang and M.F. Cohen, "Optimized color sampling for robust matting," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2007, pp. 1–8.

[8] J. Sun, J. Jia, C.K. Tang, and H.Y. Shum, "Poisson matting," *ACM Transactions on Graphics (ToG)*, vol. 23, no. 3, pp. 315–321, 2004.

[9] Y.Y. Chuang, B. Curless, D.H. Salesin, and R. Szeliski, "A bayesian approach to digital matting," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2001, vol. 2, pp. II–264.

[10] C. Rhemann, C. Rother, J. Wang, M. Gelautz, P. Kohli, and P. Rott, "A perceptually motivated online benchmark for image matting," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2009, pp. 1826–1833.

[11] A. Agrawal and Y. Xu, "Coded exposure deblurring: Optimized codes for psf estimation and invertibility," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2009, pp. 2066–2073.

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

[13] Y.W. Tai, H. Du, M.S. Brown, and S. Lin, "Image/video deblurring using a hybrid camera," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2008, pp. 1–8.

[14] Y.W. Tai, H. Du, M.S. Brown, and S. Lin, "Correction of spatially varying image and video motion blur using a hybrid camera," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 32, no. 6, pp. 1012–1028, 2010.

[15] M. McGuire, W. Matusik, and W. Yerazunis, "Practical, real-time studio matting using dual imagers," in *Proc. EGSR*, 2006, pp. 235–244.

[16] N. Joshi, W. Matusik, and S. Avidan, "Natural video matting using camera arrays," in *ACM Transactions on Graphics (TOG)*. ACM, 2006, vol. 25, pp. 779–786.

[17] Y.W. Tai, N. Kong, S. Lin, and S.Y. Shin, "Coded exposure imaging for projective motion deblurring," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2010, pp. 2408–2415.

[18] S. Dai and Y. Wu, "Removing partial blur in a single image," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2009, pp. 2544–2551.

[19] Q. Shan, W. Xiong, and J. Jia, "Rotational motion deblurring of a rigid object from a single image," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2007.

[20] Y.-W. Tai, N. Kong, S. Lin, and S. Y. Shin, "Coded exposure imaging for projective motion deblurring," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2010.

[21] S. Cho and S. Lee, "Fast motion deblurring," in *ACM Transactions on Graphics (SIGGRAPH ASIA)*, 2009.

[22] L. Xu and J. Jia, "Two-phase kernel estimation for robust motion deblurring," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2010.

[23] M. Hirsch, C. J. Schuler, S. Harmeling, and B. Schölkopf, "Fast removal of non-uniform camera-shake," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2011.

[24] M. Hirsch, S. Sra, B. Schölkopf, and S. Harmeling, "Efficient filter flow for space-variant multiframe blind deconvolution," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2010.