Image/Video Deblurring using a Hybrid Camera

Yu-Wing Tai[†*] yuwing@nus.edu.sg
Hao Du[‡*] duhao@fudan.edu.cn
Michael S. Brown[†] brown@comp.nus.edu.sg
Stephen Lin[§] stevelin@microsoft.com
National University of Singapore[†] Fudan University[‡] Microsoft Research Asia[§]

Abstract

We propose a novel approach to reduce spatially varying motion blur using a hybrid camera system that simultaneously captures high-resolution video at a low-frame rate together with low-resolution video at a high-frame rate. Our work is inspired by Ben-Ezra and Nayar [3] who introduced the hybrid camera idea for correcting global motion blur for a single still image. We broaden the scope of the problem to address spatially varying blur as well as video imagery. We also reformulate the correction process to use more information available in the hybrid camera system, as well as iteratively refine spatially varying motion extracted from the low-resolution high-speed camera. We demonstrate that our approach achieves superior results over existing work and can be extended to deblurring of moving objects.

1. Introduction

We present a novel approach to reduce spatially varying motion blur in images or video. Our work is inspired by the hybrid camera framework introduced by Ben-Ezra and Nayar [3, 4] in which a camera simultaneously captures a high-resolution image together with a sequence of low-resolution images that are temporally synchronized. In their works, optical flow is derived from the low-resolution images to compute the global motion blur of the high-resolution image. With this computed global motion kernel, deconvolution is performed to correct blur in the high-resolution image. These works clearly demonstrated that a (presumably cheap) auxiliary low-resolution device with fast temporal sampling could be coupled with a high-resolution device for global motion deblurring.

Figure 1 illustrates the tradeoff between a high resolution image captured at a low frame rate, and a low resolution image captured at a high frame rate. For comparable levels of scene exposure per pixel, the high resolution image requires a longer exposure time and thus suffers from motion blur. On the other hand, a short exposure suffices for the low resolution image captured with larger sensor units, and it is therefore sharp but lacking in detail.

(a) (b)

Figure 1. Tradeoff between resolution and frame rates. (a) Image from a high resolution, low frame rate camera. (b) Image from a low resolution, high frame rate camera.

Our paper re-visits this hybrid-camera idea, but addresses the broader problem of deblurring with spatially varying motion blur. In addition, our work aims to achieve improved deblurring performance by more comprehensively exploiting the information available in the hybrid camera system, including optical flow, back-projection constraints between low-resolution and high-resolution images, and temporal coherency along image sequences. We also introduce an iterative deblurring algorithm that refines the estimated motion kernels in the deblurring process.

The central idea in our formulation is to combine the benefits of both deconvolution and super-resolution. Deconvolution of motion blurred, high-resolution images yields high frequency details, but with ringing artifacts due to lack of low-frequency components. In contrast, super-resolution-based reconstruction from low-resolution images recovers artifact-free low-frequency results that lack high-frequency detail. We show that the deblurring information from deconvolution and super-resolution are complementary to each other, and can be used together to elevate deblurring performance. In video deblurring applications, our method furthermore capitalizes on additional deconvolution constraints that can be derived from consecutive video frames. We demonstrate that this approach produces excellent results in deblurring spatially varying motion blur. We also present an extension of this work to moving objects, and discuss possible future research directions.

*This work was done while Yu-Wing Tai and Hao Du were visiting students at Microsoft Research Asia.

978-1-4244-2243-2/08/$25.00 ©2008 IEEE

2. Related Work

Motion deblurring can be cast as the deconvolution of an image that has been convolved with either a global motion point spread function (PSF) or a spatially varying PSF. The problem is inherently ill-posed as there are a number of unblurred images that can produce the same blurred image after convolution. Nonetheless, this problem is well studied given its utility in photography and video capture. The following describes several related works.

The majority of related work lies in traditional blind deconvolution approaches that simultaneously estimate a global motion PSF and the deblurred image. These methods include well-known algorithms such as Richardson-Lucy [27, 21] and Wiener deconvolution [31]. For a survey on blind deconvolution, readers are referred to [12]. These traditional approaches often produce less than desirable results that include artifacts such as ringing.

A recent trend in motion deblurring is to either constrain the solution of the deblurred image or to use auxiliary information to aid in either the PSF estimation or the deconvolution itself (or both). Examples include work by Fergus *et al.* [10], which used natural image statistics to constrain the solution to the deconvolved image. Raskar *et al.* [25] altered the shuttering sequence of a traditional camera to be make the PSF more suitable for deconvolution. Jia [15] used an extracted alpha mask of the blurred region to aid in the PSF estimation.

Other recent approaches use more than one image to aid in the deconvolution process. Bascle *et al.* [2] used a blurry image sequence to generate a single unblurred image. Yuan *et al.* [32] used a pair of images, one noisy and one blurred. Rav-Acha and Peleg [26] consider images that have been blurred in orthogonal directions to help estimate the PSF and constrain the resulting image. Work most closely related to ours is that of Ben-Ezra and Nayar [3, 4], which used an additional imaging sensor to capture low resolution imagery for the purpose of computing optical flow and estimating a global PSF.

The approaches mentioned above assume blur arises from a global PSF. Recent works addressing spatially varying motion blur include that of Levin [18], which used image statistics to correct a single motion blur on a stable background. Bardsley *et al.* [1] segmented the image into regions exhibiting similar blur, while Cho *et al.* [7] used two blurred images to simultaneously estimate local PSFs as well as deconvolve the two images. Their approaches [18, 1, 7], however, assume the motion blur to be globally invariant within each separated layer. Work by Shan *et al.* [28] allows the PSF to be spatial varying; however, it only addressed rotational motion.

While various previous works are related in part, our work is unique in its focus on spatially varying blur with no assumption on motion path. Moreover, our approach takes full advantage of the rich information available in the hybrid camera framework to achieve sharper and cleaner results in comparison to state-of-the-art techniques, as demonstrated in Section 6.

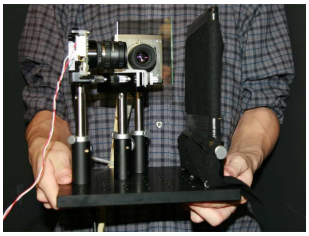


Figure 2. Our hybrid camera combines a Point Grey Dragonfly II camera, which captures images of 1024×768 resolution at 25 fps (6.25 fps for image deblurring examples), and a Mikrotron MC1311 camera that captures images of 128×96 resolution at 100 fps. A beamsplitter is employed to align their optical axes and respective images.

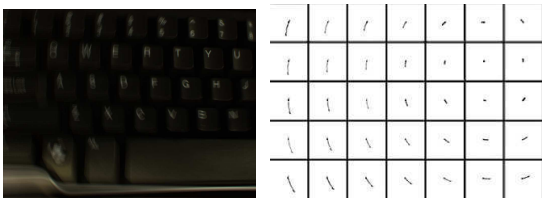
3. Hybrid Camera System

The advantages of a hybrid camera system are derived from the additional data acquired by the low-resolution high-frame-rate (LR-HFR) camera. While the spatial resolution of this camera is too low for most practical applications, the high-speed imagery is reasonably blur free and is thus suitable for optical flow computation. Since the cameras are assumed to be synchronized temporally and observing the same scene, the optical flow corresponds to the motion of the scene observed by the high-resolution low-frame-rate (HR-LFR) camera, whose images are blurred due to its slower temporal sampling. An obvious connection is to use the optical flow to compute the overall blur kernel of the high-resolution image for deconvolution.

In the following, we discuss the construction of a hybrid camera, the optical flow and motion blur estimation, and the use of the low-resolution images as reconstruction constraints on the high-resolution images.

3.1. Camera Construction

Three conceptual designs of the hybrid camera system were discussed by Ben-Ezra and Nayar [3]. In their work, they implemented a simple design in which the two cameras are placed side-by-side, such that their viewpoints can be considered the same when viewing a distant scene. A second design avoids the distant scene requirement by using a beam splitter to share between the two sensing devices the light rays that pass through a single aperture, as demonstrated by McGuire *et al.* [24] for the studio matting problem. A promising third design is to capture both the HR-LFR and LR-HFR video on a single sensor chip. According to [5], this can readily be achieved using a programmable CMOS sensing device.



(a) (b)

Figure 3. Spatially varying blur kernel estimation using optical flows. (a) Motion blur image. (b) Estimated blur kernels of (a) using optical flow.

In our work, we built a hand-held hybrid camera system based on the second design as shown in Figure 2. The two cameras were positioned such that their optical axes and pixel arrays are well aligned, and video synchronization is achieved using a 8051 microcontroller. To match the color responses of the two devices, we employ histogram equalization. Since the exposure levels of the two devices are set to be equal, the signal-to-noise ratios in the HR-LFR and LR-HFR images are approximately the same.

3.2. Blur Kernel Approximation Using Optical Flow

In the absence of occlusion, disocclusion, and out-of-plane rotation, a blur kernel can be assumed to represent the motion of a camera relative to objects in the scene. In [3], this relative motion is assumed to be constant throughout an image, and the globally invariant blur kernel is obtained through the integration of global motion vectors over a spline curve.

However, since optical flow is in fact a local estimation of motions, we can calculate spatially varying blur kernels from optical flow. We use the pyramidal Lucas-Kanade algorithm [20] to calculate the optical flow at each pixel location. These per-pixel motion vectors are then integrated to form spatially varying blur kernels, one per pixel. Figure 3 shows an example of spatially varying blur kernels estimated from optical flows.

These estimated blur kernels contain quantization errors due to the low resolution of the optical flows. Additionally, motion vector integration may provide an imprecise temporal interpolation of the flow observations. In our Bayesian optimization framework, we consider the blur kernels to also be parameters to be estimated, and thus are subject to refinement in the optimization procedure. The details will be discussed fully in Section 4.

3.3. Back-Projection Constraints

The capture of low-resolution frames in addition to the high-resolution images not only facilitates optical flow computation, but also provides super-resolution-based reconstruction constraints [13, 14, 6, 9, 29] on the high-resolution deblurring solution. The back-projection algorithm [13, 14] is one of the most common iterative techniques to minimize reconstruction error, and can be formulated as follows:

$$I^{t+1} = I^t + \sum_{j=1}^{M} (u(I_{l_j} - d(I^t \otimes h))) \otimes p \qquad (1)$$

where M is the number of corresponding low-resolution observations, t is an iteration index, I_{l_j} is the j-th low-resolution image, $\otimes$ is the convolution operation, h is the convolution filter before downsampling, p is a filter representing the back-projection process, and $d(\cdot)$ and $u(\cdot)$ are the downsampling and upsampling processes respectively. Equation (1) assumes that the low-resolution images have been aligned and that each observation carries the same weight. In the absence of a prior, h is chosen to be a gaussian filter with a size proportionate to the downsampling factor, and p is set equal to h.

In the hybrid camera system, a number of low-resolution frames are captured in conjunction with each high-resolution image. To exploit this available data, we align these frames according to the computed optical flows, and use them as back-projection constraints in Equation (1). The number of low-resolution image constraints M is determined by the relative frame rates of the cameras. Each of the low-resolution frames presents a physical constraint on the high-resolution solution, in a manner resembling a set of offset images in the super-resolution application.

4. Bayesian Optimization Framework

A brief review on the Richardson-Lucy deconvolution algorithm is given as our approach is fashioned in a similar manner. For sake of clarity, our algorithm is first discussed for use with correcting global motion blur, followed by its extension to spatially varying blur kernels.

4.1. Richardson-Lucy Image Deconvolution

The Richardson-Lucy algorithm [27, 21] is an iterative deconvolution algorithm derived from Bayes Theorem that minimizes the following estimation error:

$$\arg\min_{I} n(||I_b - I \otimes K||^2) \qquad (2)$$

where I is the deblurred image, K is the blur kernel, I_b is the observed blur image, and $n(\cdot)$ is the image noise distribution. A solution can be obtained using the iterative update algorithm defined as follows:

$$I^{t+1} = I^t \times K * \frac{I_b}{I^t \otimes K} \qquad (3)$$

where $*$ is the correlation operation. A blind deconvolution method using the Richardson-Lucy algorithm was proposed by Fish *et al.* [11], which iteratively optimizes I and K in alternation. The same equation (3) was used with positions

of I and K switched during optimization iterations for K. The Richardson-Lucy algorithm assumes image noise $n(\cdot)$ to follow a Poisson distribution. If we assume image noise $n(\cdot)$ to follow a Gaussian distribution, then a least squares method can be employed [13]:

$$I^{t+1} = I^t + K * (I_b - I^t \otimes K) \tag{4}$$

which shares the same iterative back-projection update rule as Equation (1).

From video input with computed optical flows, multiple blurred images I_b and blur kernels K may be acquired by reversing the optical flows of neighboring high-resolution frames. These multiple observation constraints can be jointly applied in Equation (4) [26] as

$$I^{t+1} = I^t + \sum_{i=1}^{N} w_i K_i * (I_{b_i} - I^t \otimes K_i) \tag{5}$$

where N is the number of aligned observations. That image restoration can be improved with additional observations under different motion blurs is a important property that we exploit in this work. The use of neighboring frames in this manner may also serve to enhance the temporal consistency of the deblurred video frames.

4.2. Optimization for Global Kernels

In solving for the deblurred images, our method jointly employs the multiple deconvolution and back-projection constraints derived from the hybrid camera input. For simplicity, we assume in this subsection that the blur kernels are spatially invariant. Our approach can be formulated into a Bayesian framework as follows:

$$\begin{aligned} &\arg\max_{I,K} P(I,K|I_b,K_o,I_l) \\ =\ &\arg\max_{I,K} P(I_b|I,K)P(K_o|I,K)P(I_l|I)P(I)P(K) \\ =\ &\arg\min_{I,K} L(I_b|I,K)+L(K_o|I,K)+L(I_l|I)+L(I)+L(K) \end{aligned} \tag{6}$$

where I and K are the sharp images and the blur kernels we want to estimate, I_b, K_o and I_l are the observed blur images, estimated blur kernels from optical flows, and the high frame rate low resolution images respectively, and $L(\cdot) = -log(P(\cdot))$. In our formulation, we make no assumption on the priors $P(I)$ and $P(K)$. Assuming that $P(K_o|I,K)$ is conditionally independent of I, that the estimation errors of likelihood probabilities $P(I_b|I,K)$, $P(K_o|K)$ and $P(I_l|I)$ follow Gaussian distributions and that each observation of I_b, K_o and I_l are independent and identically distributed, we can then rewrite Equation (6) as

$$\begin{aligned} \arg\min_{I,K} &\sum_i^N ||I_{b_i} - I \otimes K_i||^2 + \lambda_B \sum_j^M ||I_{l_j} - d(I \otimes h)||^2 \\ &+\lambda_K \sum_i^N ||K_i - K_{o_i}||^2 \end{aligned} \tag{7}$$

(a) (b) (c)

Figure 4. Convolution with kernel decomposition. (a) Convolution result without kernel decomposition, where full-sized kernels are generated on-the-fly per-pixel. (b) Convolution using 30 PCA-decomposed kernels. (d) Convolution using delta function decomposition of kernels, with at most 30 delta functions per pixel.

where λ_K and λ_B are the relative weights of the error terms. To optimize the above equation for I and K, we employ alternating minimization. Combining Equations (1) and (5) yields our iterative update rules:

1. Update $I^{t+1} = I^t + \sum_{i=1}^N K_i^t * (I_{b_i} - I^t \otimes K_i^t)$ $+\lambda_B \sum_{j=1}^M h \otimes (u(I_{l_j} - d(I^t \otimes h)))$

2. Update $K_i^{t+1} = K_i^t + \widetilde{I}^{t+1} * (I_{b_i} - I^{t+1} \otimes K_i^t)$ $+\lambda_K (K_{o_i} - K_i^t)$

where $\widetilde{I} = I / \sum_{(x,y)} I(x,y)$, $I(x,y) \geq 0$, $K_i(u,v) \geq 0$, and $\sum_{(u,v)} K_i(u,v) = 1$. The two steps are updated in alternation until the change in I falls below a specified level. In our implementation, we set $N = 3$ in correspondence to the current, previous and next frames, and M is set according to the relative camera settings (4/16 for video/image deblurring). We also initialize $I^0 = I_b$ (the currently observed blurred image), $K_i^0 = K_{o_i}$ (the estimated blur kernel from optical flows), and set $\lambda_B = \lambda_K$.

4.3. Spatially Varying Kernels

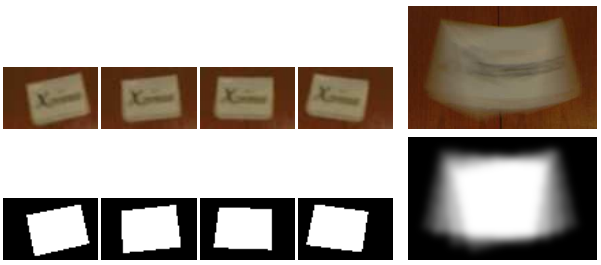
A spatially varying blur kernel can be expressed as $K(x,y,u,v)$, where (x,y) is the image coordinate and (u,v) is the kernel coordinate. For large sized kernels, e.g. 65×65, this representation is impractical due to its enormous storage requirements. Recent work has suggested ways to reduce the storage size by constraining the motion path [28]; however, our approach places no constraints on possible motion. Instead, we decompose the spatially varying kernels into a set of P basis kernels k_l whose mixture weights a_l are a function of image location:

$$K(x,y,u,v) = \sum_{l=1}^{P} a_l(x,y) k_l(u,v). \tag{8}$$

The convolution equation then becomes

$$I(x,y) \otimes K(x,y,u,v) = \sum_{l=1}^{P} a_l(x,y)(I(x,y) \otimes k_l(u,v)). \tag{9}$$

In [17], principal components analysis (PCA) is used to find the basis kernels. PCA, however, does not guarantee positive kernel values, and we have found in our experiments that PCA-decomposed kernels lead to unacceptable



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

Figure 5. Layer separation using a hybrid camera: (a)-(d) Low resolution frames and its corresponding binary segmentation masks. (e) High resolution frame and the matte estimated by compositing the low resolution segmentation masks with smoothing.

ringing artifacts, exemplified in Figure 4. We propose instead to use a delta function representation, where each delta function represents a position (u, v) within a kernel. Since a motion blur kernel is typically sparse, we store only $30 \sim 40$ delta functions for each image pixel, where the delta function positions are determined by the initial optical flows. This normally results in about $500 \sim 600$ distinct delta functions in total for an entire image, and provides a sufficient approximation of the spatially varying blur kernels in the convolution process.

Combining Equations (9) and (7), our optimization function becomes

$$\arg\min_{I,K} \sum_{i}^{N} ||I_{b_i} - \sum_{l}^{P} a_{il}(I \otimes k_{il})||^2 + \lambda_B \sum_{j}^{M} ||I_{l_j} - d(I \otimes h)||^2 + \lambda_K \sum_{i}^{N} \sum_{l}^{P} ||a_{il}k_{il} - a_{o_{il}}k_{il}||^2. \quad (10)$$

The corresponding iterative update rules are then

1. Update $I^{t+1} = I^t + \sum_{i=1}^{N} \sum_{l}^{P} a_{il}^t k_{il} * (I_{b_i} - \sum_{l}^{P} a_{il}^t (I^t \otimes k_{il})) + \lambda_B \sum_{j=1}^{M} h \otimes (u(I_{l_j} - d(I^t \otimes h)))$

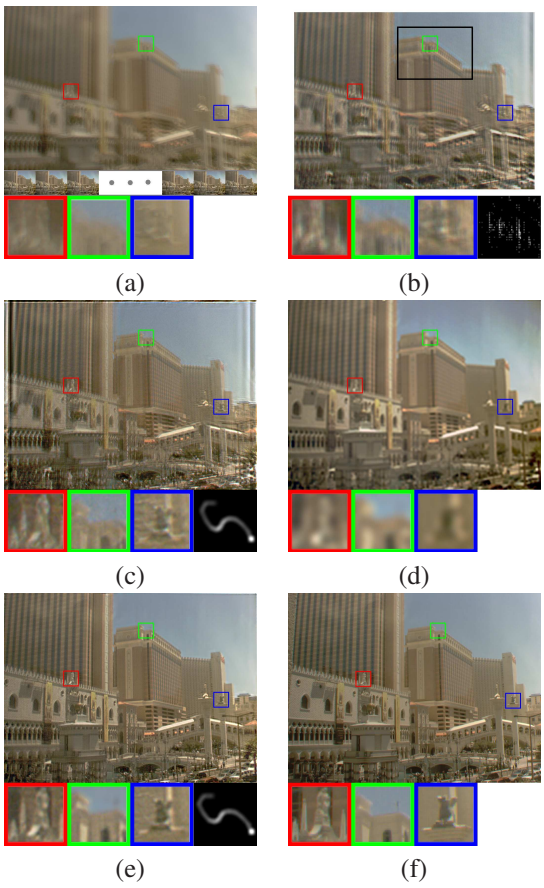
2. Update $a_{il}^{t+1} = a_{il}^t + (\widetilde{I'}^{t+1} * (I'_{b_i} - \sum_{l}^{P} a_{il}^t (I'^{t+1} \otimes k_{il}))) \cdot k_{il} + \lambda_K (a_{o_{il}} - a_{il}^t)$

where I' and I'_b are local patches of the estimated result and the blur image. The number of delta functions k_{il} stored at each pixel position may be reduced when an updated value of a_{il} becomes insignificant. Since convolution is a linear operation, each update rule itself is a convex optimization problem such that the energy in Equation (10) is monotonically non-increasing with each step. For greater stability, we process each update rule five times before switching to the other.

4.4. Discussion

Utilizing both deconvolution of high-resolution images and back-projection from low-resolution images offers distinct advantages, because the deblurring information from these two sources tend to complement each other. This can be intuitively seen by considering a low-resolution image to be a sharp high-resolution image that has undergone motion blurring with a gaussian PSF and bandlimiting. Back-projection may then viewed as a deconvolution with a gaussian blur kernel, and would promote recovery of lower-frequency image features without artifacts. On the other hand, deconvolution of high-resolution images with the high-frequency PSFs typically associated with camera and object motion generally supports reconstruction of higher-frequency details, especially those orthogonal to the motion direction. While some low-frequency content can also be restored from motion blur deconvolution, there is often significant loss due to the large support regions for motion blur kernels, and this results in ringing artifacts. As discussed in [26], the joint use of images having such different blur functions and deconvolution information favors a better deblurring solution.

Multiple motion blur deconvolutions and multiple back-



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

Figure 6. Image deblurring using globally invariant kernels. (a) Input. (b) Result generated by [10], where the user-selected regions are indicated by black boxes. (c) Result generated by [3]. (d) Result generated by back projection [13]. (e) Our results. (f) The ground truth sharp image. Close-up views and the estimated global blur kernels are also shown.

projections can further help to generate high quality results. Differences in motion blur kernels among neighboring frames provide different frequency information; and multiple back-projection constraints help to reduce quantization and the effects of noise in low-resolution images. In some circumstances there exists redundancy in information from a given source, such as when high-resolution images contain identical motion blur, or when low-resolution images are offset by integer pixel amounts. This makes it particularly important to utilize as much deblurring information as can be obtained.

Although our approach can acquire and utilize a greater amount of data, high frequency details that have been lost by both motion blur and downsampling cannot be recovered. This is a fundamental limitation of any deconvolution algorithm. We also note that reliability in optical flow cannot be assumed beyond a small time interval. This places a restriction on the number of motion blur deconvolution constraints that can be employed to deblur a given frame.

5. Extension to Deblurring of Moving Objects

In the presence of moving objects (and thus occlusion and disocclusion), the high-resolution image needs to be segmented into different layers, because pixels on the blended boundaries of moving objects contain both foreground and background components, each with different relative motion to the camera. This layer separation is inherently a matting problem which typically requires user assisted extraction [8, 19]. Fully automatic approaches, however, have required either a blue background [30], multiple cameras with different focus [23], polarized illumination [24] or a camera array [16]. In this section, we propose a simple solution to layer separation that takes advantage of the hybrid camera system. As matting is not the primary focus of our work, we offer a brief description of our basic approach to moving object layer separation, and leave a more detailed investigation of this problem for future study.

In a hybrid-camera setup, moving objects should still remain sharp in high frame rate video. To extract the alpha matte of a moving object, we can perform binary segmentation of the moving object in the low resolution images, and then compose the binary segmentation masks with smoothing to approximate the alpha matte in the high-resolution image. In Figure 5, an example of this matte extraction is demonstrated together with the moving object separation method of Zhang *et al.* [33]. Once the moving objects are separated, we can then deblur each layer separately using our framework. The alpha mattes are also deblurred for compositing, and the occluded background areas revealed after alpha mask deblurring can then be filled in either by back-projection from the low-resolution images or by the motion inpainting method of [22].

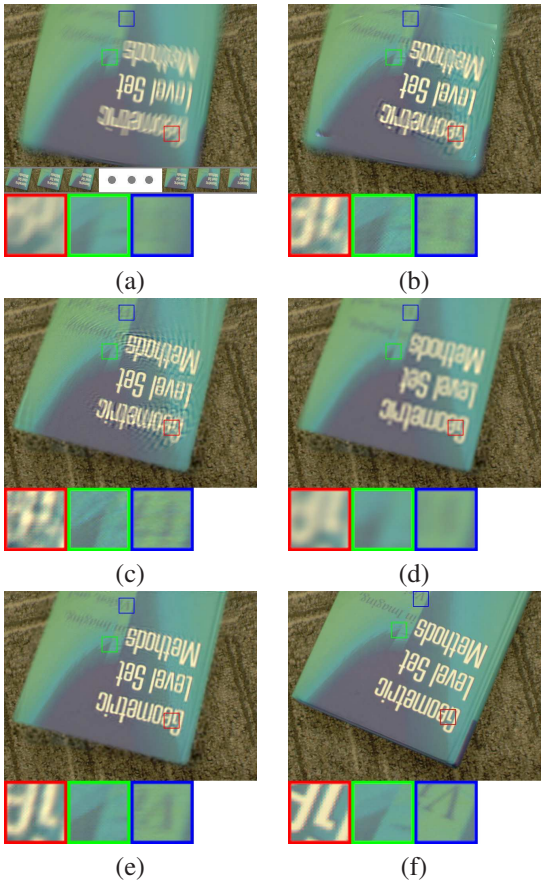


Figure 7. Image deblurring with spatial varying kernels from rotational motion. (a) Input. (b) Result generated by [28] (Result is obtained courtesy of the authors of [28]). (c) Result generated by [3] with spatially varying blur kernels estimated from optical flow. (d) Result generated by back projection [13]. (e) Our results. (f) The ground truth sharp image. Close-ups are also shown.

6. Results and Comparisons

We evaluate our approach using real images on both image deblurring and video deblurring. In these experiments, a ground-truth blur-free image is acquired by mounting the camera on a tripod and capturing a static scene. Motion blurred images are then obtained by moving the camera and/or introducing a dynamic scene object.

In Figure 6, we present an image deblurring example with globally invariant motion, where the input is one high-resolution image and several low-resolution images. Our results are compared with those generated from Fergus *et al.* [10], Ben-Ezra and Nayar [3] and back projection [13]. Fergus *et al.*'s approach is a state-of-the-art blind deconvolution technique that employs a natural image statistics constraint. However, when the blur kernel is not correctly estimated, an unsatisfactory result shown in (b) will be produced. Ben-Ezra and Nayar use the estimated optical flow as the blur kernel and then perform deconvolution. Their

result in (c) is better than that in (b) as the estimated blur kernel is more accurate, but ringing artifacts are still unavoidable. Back-projection produces a super-resolution result from a sequence of low resolution images as shown in (d). Noting that motion blur removal is not the intended application of back-projection, we can see that its results are blurry since the high-frequency details are not sufficiently captured in the low resolution images. The result of our method and the refined kernel estimate are displayed in (e). The ground truth is given in (f) for comparison.

Figure 7 shows an example with rotational motion. We compared our result with those by Shan *et al.* [28], Ben-Ezra and Nayar [3], and back projection [13]. Our approach is seen to produce less ringing artifacts compared to [28] and [3], and it generates greater detail than [13].

The benefit of using multiple deconvolutions from multiple high-resolution frames is exhibited in Figure 8, for a pinwheel with both translational and rotational motion. The deblurring result in (c) was computed using only (a) as input. Likewise, (d) is the deblurred result from only (b). Using both (a) and (b) as inputs yields the improved result in (e). This improvement can be attributed to the difference in high-frequency detail that can be recovered from each of the differently blurred images. The ground truth is shown in (f) for comparison.

Figure 9 demonstrates video deblurring of a tossed box with arbitrary motion. The top row displays five consecutive frames of input. The middle row shows our separated mattes for the moving object, and the last row presents our results. The information gained from consecutive frames leads to high-quality deblurring results and enhanced temporal consistency.

7. Conclusion

We have proposed an approach for image/video deblurring using a hybrid camera. Our work has formulated the deblurring process as an iterative method that incorporates optical flow, back-projection, kernel refinement, and frame coherency to effectively combine the benefits of both deconvolution and super-resolution. We demonstrate that this approach can produce results that are sharper and cleaner than state-of-the-art techniques.

Future research directions for this work include enlarging the temporal resolution of the high-resolution deblurred video. With the deblurring of each individual frame, temporal discontinuities in the low-frame-rate video become more evident and may result in some jumpiness through the sequence. Another research direction is to allow greater refinement of blur kernels in the iterative optimization. Currently, the refinement is limited to within the low-resolution PSF estimated from optical flows. We believe that this hybrid camera framework has significant potential for extension into other application domains.

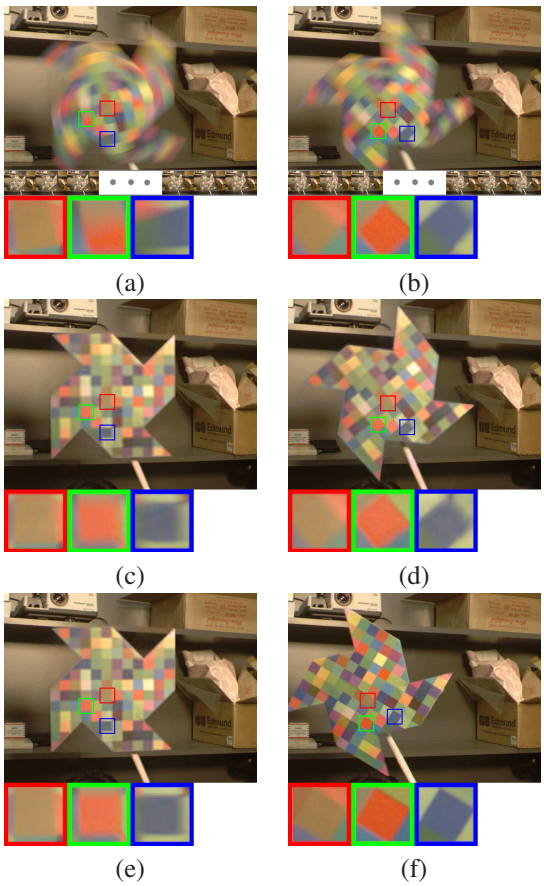


Figure 8. Deblurring with and without multiple high-resolution frames. (a)(b) Input images containing both translational and rotational motion blur. (c) Deblurring using only (a) as input. (d) Deblurring using only (b) as input. (e) Deblurring of (a) using both (a) and (b) as inputs. (f) Ground truth sharp image. Close-ups are also shown.

Acknowledgements

We thank Bennett Wilburn for providing the high speed cameras used in our implementation.

References

[1] J. Bardsley, S. Jefferies, J. Nagy, and R. Plemmons. Blind iterative restoration of images with spatially-varying blur. In *Optics Express*, pages 1767–1782, 2006.

[2] B. Bascle, A. Blake, and A. Zisserman. Motion deblurring and super-resolution from an image sequence. In *ECCV*, pages 573–582, 1996.

[3] M. Ben-Ezra and S. Nayar. Motion Deblurring using Hybrid Imaging. In *CVPR*, volume I, pages 657–664, Jun 2003.

[4] M. Ben-Ezra and S. Nayar. Motion-based Motion Deblurring. *IEEE Trans. PAMI*, 26(6):689–698, Jun 2004.

[5] M. Bigas, E. Cabruja, J. Forest, and J. Salvi. Review of cmos image sensors. *Microelectronics Journal*, 37(5):433–451, 2006.

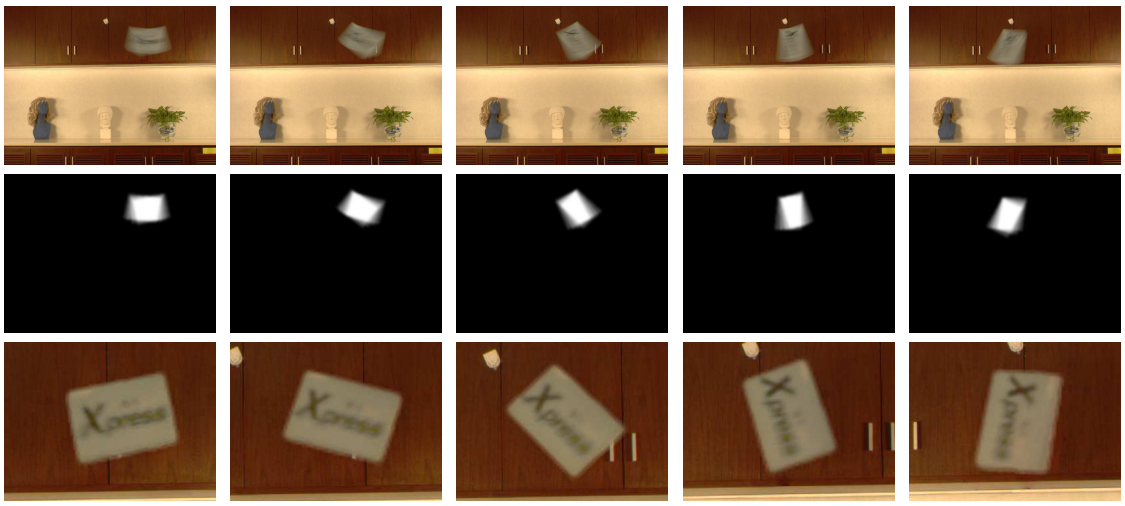


Figure 9. Video deblurring with a static background and a moving object. The moving object is a tossed box with arbitrary motion. Top: Input video frames. Middle: The alpha mattes of the moving object. Bottom: The deblurred video frames.

[6] S. Borman and R. Stevenson. Super-resolution from image sequences - a review. *mwscas*, 00:374, 1998.

[7] S. Cho, Y. Matsushita, and S. Lee. Removing non-uniform motion blur from images. In *ICCV*, 2007.

[8] Y. Chuang, B. Curless, D. H. Salesin, and R. Szeliski. A bayesian approach to digital matting. In *CVPR*, pages 264–271, 2001.

[9] M. Elad and A. Feuer. Superresolution restoration of an image sequence: adaptivefiltering approach. *IEEE Trans. Image Processing*, 8(3):387–395, 1999.

[10] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman. Removing camera shake from a single photograph. *ACM Trans. Graph.*, 25(3), 2006.

[11] D. Fish, A. Brinicombe, E. Pike, and J. Walker. Blind deconvolution by means of the richardson-lucy algorithm. *J. Opt. Soc. Am.*, 12, 1995.

[12] P. C. Hansen, J. G. Nagy, and D. P. OLeary. Deblurring images: Matrices, spectra, and filtering. *Society for Industrial and Applied Mathematic*, 2006.

[13] M. Irani and S. Peleg. Improving resolution by image registration. *CVGIP*, 53(3):231–239, 1991.

[14] M. Irani and S. Peleg. Motion analysis for image enhancement: Resolution, occlusion and transparency. *JVCIR*, 1993.

[15] J. Jia. Single image motion deblurring using transparency. In *CVPR*, 2007.

[16] N. Joshi, W. Matusik, and S. Avidan. Natural video matting using camera arrays. *ACM Trans. Graph.*, 2006.

[17] T. Lauer. Deconvolution with a spatially-variant psf. In *Astr. Data Analysis II*, volume 4847, pages 167–173, 2002.

[18] A. Levin. Blind motion deblurring using image statistics. In *NIPS*, pages 841–848, 2006.

[19] A. Levin, D. Lischinski, and Y. Weiss. A closed form solution to natural image matting. In *CVPR*, 2006.

[20] B. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *Proceedings of Imaging understanding workshop*, pages 121–130, 1981.

[21] L. Lucy. An iterative technique for the rectification of observed distributions. *Astron. J.*, 79, 1974.

[22] Y. Matsushita, E. Ofek, W. Ge, X. Tang, and H. Shum. Full-frame video stabilization with motion inpainting. *IEEE Trans. PAMI*, 28(7), 2006.

[23] M. McGuire, W. Matusik, H. Pfister, J. F. Hughes, and F. Durand. Defocus video matting. *ACM Trans. Graph.*, pages 567–576, 2005.

[24] M. McGuire, W. Matusik, and W. Yerazunis. Practical, real-time studio matting using dual imagers. In *EGSR*, 2006.

[25] R. Raskar, A. Agrawal, and J. Tumblin. Coded exposure photography: motion deblurring using fluttered shutter. *ACM Trans. Graph.*, 25(3), 2006.

[26] A. Rav-Acha and S. Peleg. Two motion blurred images are better than one. *PRL*, 26:311–317, 2005.

[27] W. Richardson. Bayesian-based iterative method of image restoration. *J. Opt. Soc. Am.*, 62(1), 1972.

[28] Q. Shan, W. Xiong, and J. Jia. Rotational motion deblurring of a rigid object from a single image. In *ICCV*, 2007.

[29] E. Shechtman, Y. Caspi, and M. Irani. Space-time super-resolution. *IEEE Trans. PAMI*, 27(4):531–544, 2005.

[30] A. Smith and J. F. Blinn. Blue screen matting. *SIGGRAPH*, 1996.

[31] Wiener and Norbert. Extrapolation, interpolation, and smoothing of stationary time series. *New York: Wiley*, 1949.

[32] L. Yuan, J. Sun, L. Quan, and H. Shum. Image deblurring with blurred/noisy image pairs. In *ACM Trans. Graph.*, page 1, 2007.

[33] G. Zhang, J. Jia, W. Xiong, T. Wong, P. Heng, and H. Bao. Moving object extraction with a hand-held camera. In *ICCV*, 2007.