Camera Intrinsic Blur Kernel Estimation: A Reliable Framework

Ali Mosleh[1] Paul Green[2] Emmanuel Onzon[2] Isabelle Begin[2] J.M. Pierre Langlois[1]

[1]École Polytechnique de Montreál, Montréal, QC, Canada [2]Algolux Inc., Montreál, QC, Canada

{ali.mosleh,pierre.langlois}@polymtl.ca

{paul.green,emmanuel.onzon,isabelle.begin}@algolux.com

Abstract

This paper presents a reliable non-blind method to measure intrinsic lens blur. We first introduce an accurate camera-scene alignment framework that avoids erroneous homography estimation and camera tone curve estimation. This alignment is used to generate a sharp correspondence of a target pattern captured by the camera. Second, we introduce a Point Spread Function (PSF) estimation approach where information about the frequency spectrum of the target image is taken into account. As a result of these steps and the ability to use multiple target images in this framework, we achieve a PSF estimation method robust against noise and suitable for mobile devices. Experimental results show that the proposed method results in PSFs with more than 10 dB higher accuracy in noisy conditions compared with the PSFs generated using state-of-the-art techniques.

1. Introduction

The quality of images formed by lenses is limited by the blur generated during the exposure. Blur most often occurs on out of focus objects or due to camera motion. While these kinds of blur can be prevented by adequate photography skills, there is a permanent intrinsic blur caused by the optics of image formation *e.g.* lens aberration and light diffraction. Image deconvolution can reduce this intrinsic blur if the lens PSF is precisely known. The PSF can be measured directly using laser and precision collimator or pinhole image analysis. However, these approaches require sophisticated and expensive equipment. Modeling the PSF by means of camera lens prescription [19] or parameterized techniques [21] is also possible. However, these techniques are often applicable only for certain camera configurations and need fundamental adjustments for various configurations. Hence, there is a requirement to measure the blur function by analyzing the captured images. Such a PSF estimation is an ill-posed problem that can be approached by blind and non-blind methods. This problem is even more challenging for mobile devices since they have very small sensor area that typically creates a large amount of noise.

Blind PSF estimation is performed on a single observed image [4, 8, 9, 11, 14, 17, 23, 25, 28] or a set of observed images [6, 16, 27]. The features of the latent sharp image are modeled, and then the model is employed in an optimization process to estimate a PSF. Given the knowledge that the gradient of sharp images generally follows a heavy-tailed distribution [20], Gaussian [26], Laplacian [3], and hyper-Laplacian [15] priors over image derivatives are used in many techniques such as [14, 8, 18, 13]. In addition to these general priors, local edges and a Gaussian prior on the PSF are used in edge-based PSF estimation techniques [4, 5, 11, 25]. In general, blind PSF estimation methods are suitable to measure the extrinsic camera blur function rather than the intrinsic one.

Non-blind PSF estimation techniques assume that given a known target and its captured image the lens PSF can be accurately estimated. Zandhuis *et al.* [29] propose to use slanted edges in the calibration pattern. Several one dimensional responses are required that are based on a symmetry assumption for the kernel. A checkerboard pattern is used as the calibration target by Trimeche *et al.* in [24], and the PSF is estimated by inverse filtering given the sharp checkerboard pattern and its photograph. Joshi *et al.*'s non-blind PSF estimation [11] relies on an arc-shaped checkerboard-like pattern. The PSF is estimated by introducing a penalty term on its gradient's norm. In a similar scheme, Heide *et al.* estimate the PSF using the norm of PSF's gradient in the optimization process [10]. They propose to use a white-noise pattern rather than regular checkerboard image or Joshi's arc-shaped pattern as the calibration target. This method also constrains the energy of the PSF by introducing a normalization prior to the PSF estimation function. Kee *et al.* propose a test chart that consists of a checkerboard pattern with complement black and white circles in each block [12]. The PSF estimation problem is solved using least squares minimization and thresholding out negative values generated in the result. A random noise target is also used in Brauers *et al.*'s PSF estimation technique [1].
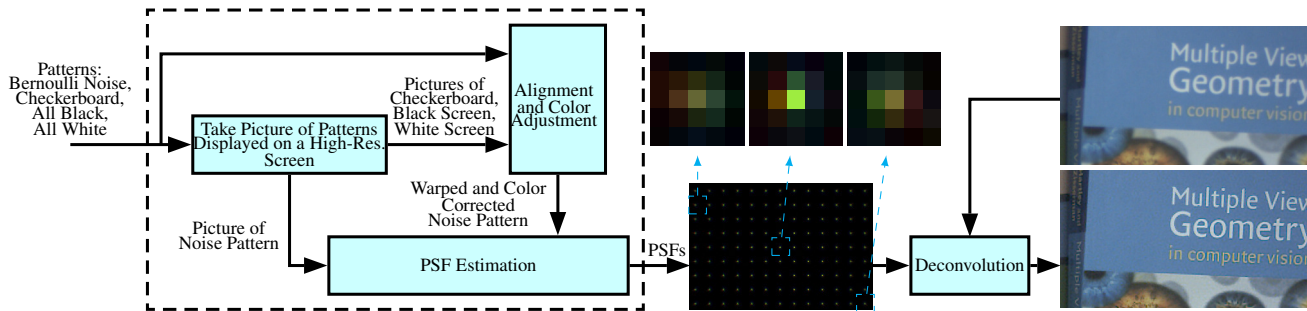
Figure 1. The overview of our lens PSF measurement framework and the enhancement achieved using our measured PSFs.

They propose to apply inverse filtering to measure the PSF, and then threshold it as a naive regularization. Delbracio *et al.* show in [7] that a noise pattern with a Bernoulli distribution with an expectation of 0.5 is an ideal calibration pattern in terms of well-posedness of the PSF estimation functional. In other words, pseudo-inverse filtering without any regularization term would result in a near optimal PSF. The downside of the direct pseudo-inverse filtering is that it does not consider the non-negativity constraint of the PSF. Hence, the PSF can be wrongly measured in presence of even a little amount of noise in the captured image. These techniques rely strongly on an accurate alignment (geometric and radiometric) between the calibration pattern and its observation. Reducing alignment errors is essential to produce accurate PSFs using these techniques.

In this paper, we introduce a non-blind method to measure the intrinsic camera blur. We build a reliable hardware setup that unlike existing non-blind techniques omits homography and radial distortion estimation for the camera-scene alignment. Hence, potential errors of the geometric alignment between the captured pattern and the original one are greatly reduced. This setup also provides pixel to pixel intensity correspondence between the captured pattern and the sharp pattern. Hence, there is no need for tone curve estimation or complicated radiometric correction between the two images. We use Bernoulli (0.5) noise patterns to estimate the PSF. Unlike the method proposed in [6], we introduce a non-negativity constraint and take into account the frequency and energy specifications of the Bernoulli noise pattern directly in the functional of the PSF estimation. Also, the proposed alignment allows us to utilize multiple PSF estimation targets (*i.e.* Bernoulli noise patterns) in the PSF estimation function to significantly reduce the effect of noise. As a result of our main contributions *i.e.* simplified and accurate alignment, employing spectral information of the kernel as a prior, and using multiple targets, we achieve an accurate PSF estimation which is greatly robust against noise. This becomes an appropriate scheme to measure lens blur of mobile devices that suffer from a large amount of noise caused by their small sensors. The accuracy of our PSF estimation method is validated by comparing with state-of-the-art non-blind PSF estimation techniques, and by deblurring images using PSFs that we measured for camera lenses.

2. Overview

Typically, a perspective projection of a 3D world scene onto a focal plane is the base of camera model. Light rays are concentrated via a system of lenses toward the focal plane passing through the aperture. It is often assumed that the observed scene i is planar. Hence, the perspective projection can be modelled as a planar homography h. The perspective projection is followed by some distortion due to the physics of imaging, especially the use of a non-pinhole aperture in real cameras. Denoting the geometric distortion function by d, image formation can be modeled as:

$$b = S\Big(v\big(d\big(h(i)\big)\big) * k\Big) + n,\qquad(1)$$

where b is the captured image, k is a PSF that represents lens aberrations, v denotes optical vignetting often caused by physical dimensions of multi-element lens, S is the sensor's sampling function, and n represents additive zero-mean Gaussian noise. It is assumed that the camera response function is linear, and for brevity, avoided in Eq. (1).

Measuring the intrinsic blur kernel k given the observed image b and a known scene i requires an accurate estimation of h, d, and v in Eq. (1). The homography h is often estimated [1, 7, 10, 11, 12, 24] using some known feature points in i (*e.g.* corners in checkerboard calibration pattern) and fitting them to the corresponding points in the observed image b, and then the effect of distortion d is taken into account by Brown's radial-tangential model [2]. After warping i according to h and d, devignetting/color correction algorithms are applied to estimate v in order to generate a sharp correspondence $u = v\big(d\big(h(i)\big)\big)$ of the observed image b to be used in the imaging model

$$b = S\left(u * k\right) + n.\qquad(2)$$

Observation-scene alignment (h, d and v estimation) is prone to severe errors. Even advanced calibration and warping techniques may negatively affect the accuracy of PSF estimation [22]. Hence, we propose to avoid traditional homography, distortion, and vignetting estimation. An overview of our PSF measurement method is shown in Fig. 1. We use four different patterns; a 0.5 expectation Bernoulli noise pattern as the scene pattern, a checkerboard with a large number of checker patterns as the calibration
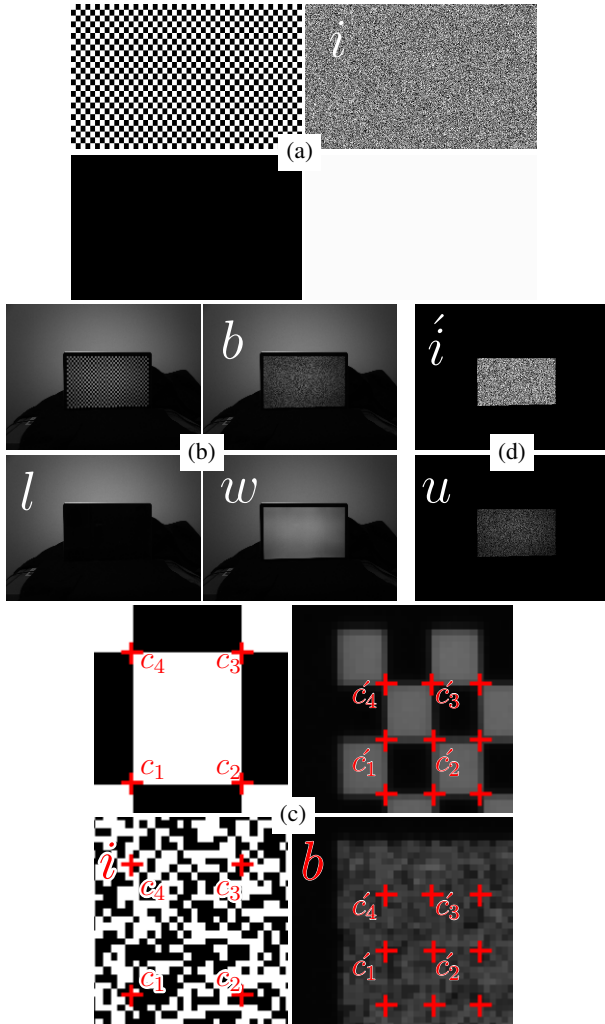
Figure 2. Patterns used in calibration and PSF estimation. (a) Original synthetic patterns. (b) Photographs of the synthetic patterns displayed on a screen. (c) Detected corners in the checkerboard images and the corresponding points in the noise images. (d) Warped and color corrected sharp noise pattern.

guide, and a black and a white image as intensity references. A high resolution screen is used to display these patterns so that no relative motion between them and between the camera and the scene is induced during the imaging. The corners found in the picture of the checkerboard are used to find the correspondence between the camera grid and the scene. These points are used in a bilinear interpolation scheme to transform the synthetic noise pattern into the camera grid space. Next, the pictures of the black and the white images are used to adjust the intensity of the transformed synthetic noise pattern. This process is further detailed in Sec. 3.1. The resulting warped and color adjusted sharp noise pattern u is then employed in our PSF estimation procedure.

Considering model (2), the lens PSF k is estimated by

generating a linear system to solve a least squares problem with smoothness and sparsity constraints for the kernel. In addition, since the spectrum of the Bernoulli pattern is uniform and contains all frequency components, we employ its spectral density function (SDF) to derive a prior for the PSF as detailed in Sec. 3.2. With this framework we can employ multiple noise patterns in order to measure the lens PSF more accurately.

3. Measuring Lens Blur

3.1. Alignment

Separating the calibration pattern from the scene i provides us with more flexibility in the size of checker blocks and the number of feature points in the calibration pattern. Fig. 2(a) shows the synthetic patterns; a 25×48 checkerboard pattern, a Bernoulli (0.5) noise pattern, a black image and a white image. The size of all of these images is chosen so that they fit the entire screen when displayed on a high resolution screen. Then, pictures of the displayed synthetic images are captured as shown in Fig. 2(b) using the camera whose lens PSF needs to be measured.

In the first step, the corner points in the pictured checkerboard and the synthetic one are detected using a Harris corner detector. By inspection, the corresponding pairs of corner points in these two images are identified. These points are in fact mapped from the synthetic sharp pattern to the camera grid through the imaging process while some lens blur is induced. Since, the geometry alignment between camera and display is unchanged between captures, the points detected in the checkerboards (Fig. 2(c)) are used to warp the sharp Bernoulli noise pattern i to align it with its corresponding captured picture b.

We denote the planar coordinates of each block identified using corner detection by $c_1 = (\alpha_1, \beta_1)$, $c_2 = (\alpha_2, \beta_1)$, $c_3 = (\alpha_2, \beta_2)$, $c_4 = (\alpha_1, \beta_2)$ in the synthetic checkerboard, and by $\acute{c}_1 = (x_1, y_1)$, $\acute{c}_2 = (x_2, y_2)$, $\acute{c}_3 = (x_3, y_3)$, $\acute{c}_4 = (x_4, y_4)$ in the pictured checkerboard (Fig. 2(c)). The synthetic noise pixels that lie in the block denoted by c_1, c_2, c_3, c_4 are mapped to the corresponding block coordinated by $\acute{c}_1$, $\acute{c}_2$, $\acute{c}_3$, $\acute{c}_4$. This is carried out by bilinear interpolation. In fact, the warping procedure can be reduced to a texture mapping from a square space into an irregular quadrilateral:

$$
\begin{pmatrix} x & y \end{pmatrix} = \begin{pmatrix} \alpha\beta & \alpha & \beta & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 & -1 & 1 \\ -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \acute{c}_1 \\ \acute{c}_2 \\ \acute{c}_4 \\ \acute{c}_3 \end{pmatrix}
$$
(3)

where (α, β) is the pixel coordinate in the square c_1, c_2, c_3, c_4. In Eq. (3), (α, β) is normalized by mapping the range $[\alpha_1, \alpha_2]$ to $[0, 1]$ and $[\beta_1, \beta_2]$ to $[0, 1]$. The transformed coordinate into the area $\acute{c}_1$, $\acute{c}_2$, $\acute{c}_3$, $\acute{c}_4$ is denoted by (x, y). For better accuracy, the pixels in the synthetic noise pattern i are

Algorithm 1 Bilinear warping.
***
Require: c_1, c_2, c_3, c_4 and $\acute{c}_1$, $\acute{c}_2$, $\acute{c}_3$, $\acute{c}_4$ for all N_{cb}
    checkerboard blocks, captured noise pattern b, syn-
    thetic noise pattern i
1: Generate $M \times N$ matrices of zeros $count$ and $\acute{i}$
2: **for all** N_{cb} blocks **do**
3:    map $[\alpha_1, \alpha_2]$ to $[0, 1]$ and $[\beta_1, \beta_2]$ to $[0, 1]$
4:    **for** $\alpha = \alpha_1$ to α_2, step: S_p **do**
5:      **for** $\beta = \beta_1$ to β_2, step: S_p **do**
6:       find x and y using Eq. (3)
7:       $count(x, y) \leftarrow count(x, y) + 1$
8:       $\acute{i}(x, y) \leftarrow \left(\acute{i}(x, y) + i(\alpha, \beta) \right) / count(x, y)$
9:      **end for**
10:    **end for**
11: **end for**
12: **return** $\acute{i}$
***

divided into S_p sub-pixels. Hence, more samples are taken into account in the warping. Assuming that N_{cb} blocks exist in the checkerboard pattern and that the size of b is $M \times N$, Algorithm 1 lists the steps to warp the synthetic noise pattern i and generate $\acute{i}$. In this algorithm, $count$ is used to keep track of pixels that are mapped from i space into a single location in the b space. This avoids rasterizarion artifacts especially at the borders of warped blocks.

The camera's vignetting effect can be reproduced by means of the pictures of black and white images *i.e.* l and w (Fig. 2(b)). Assuming that the pixel intensity ranges from 0 to 1 in $\acute{i}$, the intensity of sharp version u of the scene captured by the camera is calculated as:

$$u(x, y) = l(x, y) + \acute{i}(x, y)\big(w(x, y) - l(x, y)\big), \quad (4)$$

where $w(x, y)$ and $l(x, y)$ denote pixel intensities at (x, y) in the white and the black images (Fig. 2(b)) respectively.

Fig. 2(d) shows the result of the alignment process. Our alignment scheme avoids the estimation of the homography, distortion, and vignetting functions generally performed in state-of-the-art non-blind PSF estimation techniques. Due to the separation of calibration and target patterns, we are able to increase the number of checker patterns in the calibration image, and thus increase the accuracy of the bilinear interpolation done in the warping scheme. Our accurate vignetting reproduction is due to the use of camera reference intensities (black and white reference images), which is only possible if there is no change in the camera-scene geometry alignment while capturing the images. This in turn becomes possible by using a high-resolution screen to expose the sequence of images.

3.2. PSF estimation

The Bernoulli (0.5) noise pattern that we use in PSF estimation contains all frequency components and its spectrum does not contain zero magnitude frequencies. Therefore, it is ideal for direct estimation of PSF from b and u via inverse filtering [1, 7]. However, the presence of unknown noise in the observation b violates the expected uniform frequency in b. Hence, direct methods result in artifacts and negative values in the estimated PSF. This motivates utilizing priors in the PSF estimation.

Let $M \times N$ be the size of b and u and $R \times R$ be the size of k. Hereafter, by b and u we mean the rectangular regions in these images that contain the noise pattern. The blur model (2) can be rewritten in vector form,

$$\mathbf{b} = \mathbf{u}\mathbf{k} + \mathbf{n} \qquad (5)$$

where $\mathbf{b} \in \mathbb{R}^{MN}$, $\mathbf{n} \in \mathbb{R}^{MN}$, $\mathbf{k} \in \mathbb{R}^{RR}$, and $\mathbf{u} \in \mathbb{R}^{MN \times RR}$. For brevity, the sampling operator S is dropped as it is a linear operator that can be easily determined by measuring the pixel ratio between the synthetic image and the corresponding captured image.

The Bernoulli noise pattern has a homogeneous spectrum density function (SDF) *i.e.* $|\mathcal{F}(i)|^2$ where $\mathcal{F}(.)$ denotes the Fourier transform. Hence, in an ideal noise-free image acquisition, the SDF of the captured image b is $|\mathcal{F}(i)|^2|\mathcal{F}(k)|^2$. Therefore, the SDF of the ideal blur kernel $\acute{k}$ is expected to be

$$|\mathcal{F}(\acute{k})|^2 = \frac{\mathcal{F}(b)\overline{\mathcal{F}(b)}}{\mathcal{F}(u)\overline{\mathcal{F}(u)}}, \qquad (6)$$

where $\overline{a}$ denotes the complex conjugate of a.

We propose to solve the following function to estimate the PSF:

$$\begin{aligned}
\underset{\mathbf{k}}{\text{minimize}}\; E(\mathbf{k}) = &||\hat{\mathbf{u}}\mathbf{k} - \hat{\mathbf{b}}||^2 + \lambda||\mathbf{k}||^2 + \mu||\nabla\mathbf{k}||^2 \\
&+ \gamma|||\mathcal{F}(\mathbf{k})| - |\mathcal{F}(\acute{\mathbf{k}})|||^2, \text{ s.t. } \mathbf{k} \geqslant 0
\end{aligned}$$
$$(7)$$

where the first term is the data fitting term, and the second and the third terms are the kernel sparsity and the kernel smoothness constraints weighted by λ and μ, respectively. The last term in Eq. (7) weighted by γ is the constraint of the SDF of the PSF. Note that $||.||$ is the ℓ_2 norm and ∇ is the gradient operator. Due to the use of a screen to display the target patterns and a fixed configuration for the camera, we are able to have multiple noise patterns and their observations. Using multiple observations and sharp correspondences in problem (7) results in a more accurate PSF. In problem (7), $\hat{\mathbf{u}}$ contains L stacked number of different $\mathbf{u}$ *i.e.* $\hat{\mathbf{u}} = [\mathbf{u}_1\mathbf{u}_2\cdots\mathbf{u}_L]^T$, $\hat{\mathbf{u}} \in \mathbb{R}^{MNL \times RR}$. Similarly, $\hat{\mathbf{b}} = [\mathbf{b}_1\mathbf{b}_2\cdots\mathbf{b}_L]^T$, $\hat{\mathbf{b}} \in \mathbb{R}^{MNL}$. $|\mathcal{F}(\acute{\mathbf{k}})|$ is also calculated using multiple sharp and observation images ($\hat{\mathbf{u}}$ and $\hat{\mathbf{b}}$).

The objective function of problem (7) can be written as:

$$\begin{aligned}
E(\mathbf{k}) = &\frac{1}{2}(\hat{\mathbf{u}}^T\hat{\mathbf{u}} + \mu\mathbf{d_x}\mathbf{d_x}^T + \mu\mathbf{d_y}\mathbf{d_y}^T + \lambda)\mathbf{k}\mathbf{k}^T - \hat{\mathbf{u}}^T\mathbf{b}\mathbf{k} \\
&+ \gamma|||\mathcal{F}(\mathbf{k})| - |\mathcal{F}(\acute{\mathbf{k}})|||^2,
\end{aligned}$$
$$(8)$$

where $d_x = [-1\ 1]$ and $d_y = [-1\ 1]^T$ are the first order derivative operators whose 2D convolution vector format in Eq. (8) are $\mathbf{d_x}$ ($\mathbf{d_x} \in \mathbb{R}^{RR \times RR}$) and $\mathbf{d_y}$ ($\mathbf{d_y} \in \mathbb{R}^{RR \times RR}$) respectively. The data fitting term and the two regularization terms in Eq. (8) follow a quadratic expression whose gradient is straightforward to find. Then, the gradient of the SDF constraint in Eq. (8) can be derived as:

$$\frac{\partial|||\mathcal{F}(\mathbf{k})| - |\mathcal{F}(\acute{\mathbf{k}})|||^2}{\partial \mathbf{k}} = 2(\mathbf{k} - \mathcal{F}^{-1}\left(\frac{\mathcal{F}(b)\overline{\mathcal{F}(b)}}{\mathcal{F}(u)\overline{\mathcal{F}(u)}}e^{j\theta}\right),$$
(9)

where θ is the phase of the Fourier transform of $\acute{k}$ (Eq. (6)). We solve problem (8) by a gradient descent solver with the descent direction as $-\partial E(\mathbf{k})/\partial \mathbf{k}$.

Since the intrinsic lens blur is spatially varying, the observation and sharp images are divided into smaller corresponding blocks, and then the PSF estimation problem (7) is solved for each block independently.

4. Experimental Results

We tested the accuracy of our alignment (calibration) technique and the proposed PSF estimation method independently. Then, the entire lens PSF measurement procedure was applied on real devices and the produced PSFs were used to enhance the quality of images captured by these devices. In our experiments, an Apple Retina display with resolution 2880×1800 was used to display the patterns. Our technique was compared with state-of-the-art non-blind PSF estimation methods as detailed below.

4.1. Alignment Evaluation

We used a Ximea Vision Camera sensor MQ022CG-CM with a 12 mm lens in order to test the alignment. This lens-camera configuration was chosen as it generates a reasonable geometric and radiometric distortion. The acquisition was set so that only raw images were generated and no further process was done by the camera. The image acquisition and alignment method discussed in Sec. 3.1 was performed using the pictures of the calibration pattern and the noise target. The camera's aperture was set to be very small so that the effect of the lens blur was minimal. Images were captured in different exposure times *i.e.* 10^{-6}, 10^{-3} and 10^{-1} second, to have images with different induced noise levels. The similarity of the warped and color corrected synthetic noise pattern generated in each test was compared with the captured image using PSNR listed in Table 1. Although there is some blur in the images, the PSNR can still show the similarity between the warped synthetic pattern and the one captured by the camera.

Using the same camera-lens configuration, the geometric and radiometric calibration techniques and the calibration patterns used in [7, 11, 12] were employed to produce sharp correspondence of the captured targets. The PSNR values
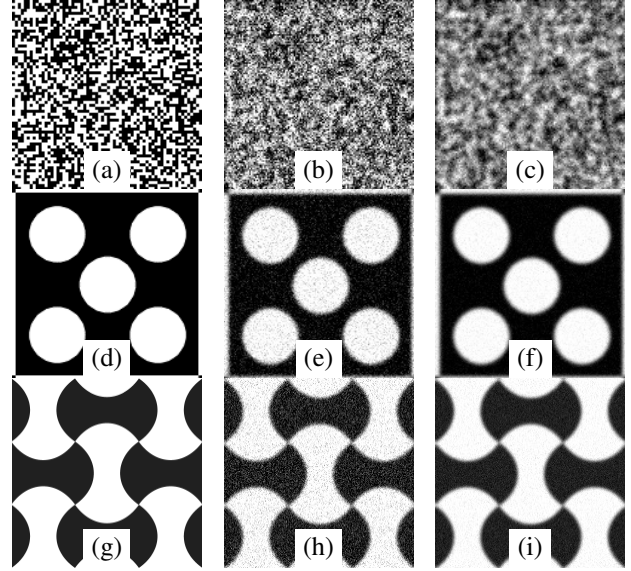

Figure 3. Synthetic data used in evaluation of the PSF estimation. (a) Our sharp Bernoulli (0.5) noise pattern. (d) Kee *et al.*'s [12] pattern. (g) Joshi *et al.*'s [11] pattern. (b,e,h) Blurred images with noise $n = \mathcal{N}(0, 0.1)$. (c,f,i) Blurred image with noise $n = \mathcal{N}(0, 0.01)$.

obtained for these results are listed in Table 1. Compared to our method, the calibration strategies used in these methods produce less accurate correspondences. The reason our technique outperforms the other methods is mainly due to the use of a display that allows us to separate the calibration pattern from the kernel estimation target. This leads to an accurate bilinear mapping since a calibration pattern with a large number of feature points (corners) can be used. Moreover, the availability of a large number of corresponding feature points helps avoid error-prone homography and distortion estimation steps. In addition, the use of a screen to display the patterns provides us with an accurate pixel to pixel intensity reference used in reproducing the camera's vignetting effect.

4.2. PSF Estimation Evaluation

Our PSF estimation using Bernoulli noise patterns was evaluated in alignment-free tests to gain an insight into its

Table 1. PSNR values in dB obtained between the warped and color corrected target and the observation (captured image of the target) using different methods.

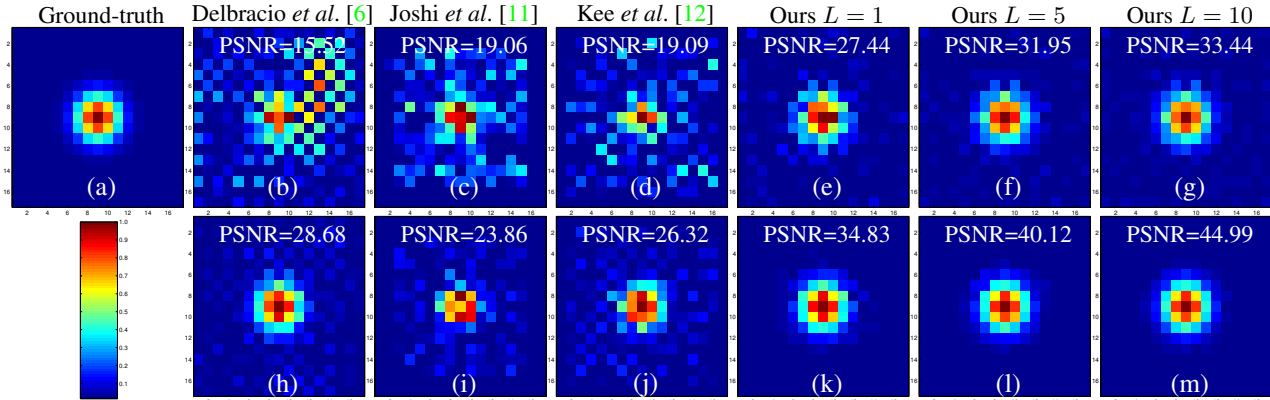| Method	Exposure (s):	10^{-6}	10^{-3}	10^{-1}
Ours		31.21	30.88	29.45
Joshi's [11]		21.30	19.64	18.98
Kee's [12]		22.42	19.36	19.05
Delbracio's [7]		22.82	20.21	18.91

Figure 4. Estimated PSFs using different non-blind techniques and their PSNRs in dB. (a) Ground-truth PSF. (b-g) Estimated PSFs in presence of noise $n = \mathcal{N}(0, 0.1)$ in b. (h-m) Estimated PSFs in presence of noise $n = \mathcal{N}(0, 0.01)$ in b.

accuracy. A sharp noise pattern was blurred according to Eq. (2). A synthetic 17×17 Gaussian kernel with standard deviation 1.5 was generated shown in Fig. 4(a) and convolved with the noise pattern. Then, zero-mean Gaussian noise n was added. Fig. 3(b) and (c) show two Bernoulli patterns blurred using the PSF shown in Fig. 4(a). The noise standard deviation is 0.1 and 0.01 in Fig. 3(b) and (c), respectively. The PSF estimation was performed given the blurry and sharp noise patterns. We set the regularization weights as $\mu = 10$, $\lambda = 10$, and $\gamma = 100$ in problem (7). Fig. 4(e) shows the estimated PSF using images shown in Fig. 3(a) and (b) and its calculated PSNR with regard to the ground-truth PSF (Fig. 4(a)). The noise corrupted the blurry image so that there is little similarity between the blurry and the sharp image. However, the estimated PSF is very similar to the ground-truth PSF (Fig. 4(a)). The PSF can be more accurately estimated by using more than one noise pattern (L factor in generating $\hat{\mathbf{u}}$ and $\hat{\mathbf{b}}$ in Eq. (7) and (8)). The resulting PSFs by choosing $L = 5$ and $L = 10$ different Bernoulli (0.5) noise patterns and their corresponding observations are illustrated in Fig. 4(f) and (g). As the number of patterns increases, the estimated PSF looks more similar to the ground-truth. It is illustrated by the obtained PSNRs. A similar test was performed on the blurry images with a lower noise level (Fig. 3(c)). Although the noise level is still considerable, the resulting PSFs (Fig. 4(k), (l) and (m)) are estimated quite accurately compared to the ground-truth PSF Fig. 4(a).

In order to gain an insight into the effect of our proposed SDF prior in PSF estimation, we performed a similar experiment with similar values for μ and λ, but with different values for γ. This time we only used one single noise pattern ($L = 1$). The noise pattern shown in Fig. 3(a) and its blurred and noisy observations were used (Fig. 3(b) and (c)). Resulting PSFs by setting the weight of the SDF prior to 0, 10 and 100 are presented in Fig. 5. As the PSNR values indicate, employing the SDF prior increases the accuracy of the PSF even though the observations (b) are very noisy.
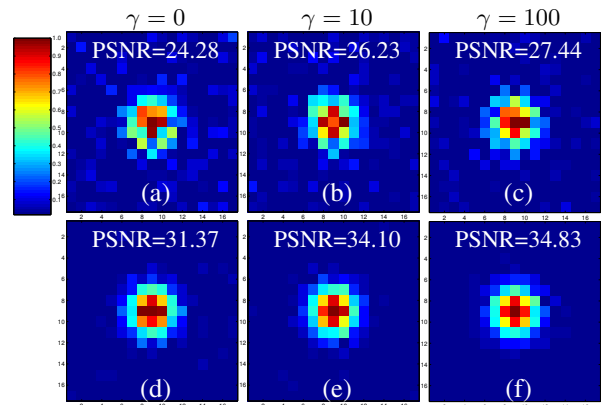

Figure 5. Effect of SDF prior in our PSF estimation (a-c) Estimated PSFs in presence of noise $n = \mathcal{N}(0, 0.1)$ in b. (d-f) Estimated PSFs in presence of noise $n = \mathcal{N}(0, 0.01)$ in b.

We estimated the PSF using Delbracio et al.'s method [7] designed to perform well on Bernoulli noise patterns. This method fails to estimate the PSF for the image that contains a noise level of 0.1 (Fig. 4(b)). Even for a lower noise level (0.01), it generates a considerable amount of artifacts in the estimated PSF (Fig. 4(h)). This occurs in the presence of even a little amount of noise, mainly due to avoiding regularization and non-negativity constraint of the PSF in the process. We simulated the same blur and noise levels on the PSF estimation targets of Joshi et al. [11] and Kee et al. [12] shown in Fig. 3(d) and (g), and then employed their proposed methods to estimate the PSF. In all cases, the proposed PSF estimation technique generates more accurate PSFs than these methods as illustrated in Fig. 4.

4.3. Experiments with Real Devices

We selected two camera devices to test the proposed PSF measurement technique; a Ximea Vision Camera (MQ022CG-CM) sensor whose resolution is 2048×1088 with a 12 mm lens, and a Blackberry mobile phone's front facing camera with resolution 1600×1200. Unlike SLR cameras, these cameras have small pixel sensors and create

a large amount of noise. Hence, it is more challenging to measure their lens blur. Camera-target alignment was performed as explained in Sec. 3.1. The checkerboard pattern and the white and black patterns Fig. 2(a) were used in the alignment, and 5 different Bernoulli noise patterns ($L = 5$) were used in the PSF estimation. The image acquisition was done in RAW format, so that PSF measurement was performed for each of the 4 different color channels that exist in the Bayer's grid. This avoids demosaicing, white-balancing, and any other post/pre-processing typically done in cameras. It is critical not to estimate a single PSF for all the 4 channels as this results in chromatic aberrations once used in a deconvolution [10]. Since the PSFs vary spatially in the camera space, PSF estimation was carried out on non-overlapping blocks of 128×128. The distance between the camera and the display was set to maintain a 1:2 ratio between the camera pixels and the screen pixels (S in Eq. (1) and (2)). Note that the screen may not cover the whole camera grid (*e.g.* Fig. 2(b)). Therefore, the whole process should be performed for various placements of the display until the PSFs are estimated for the entire camera grid. For both cameras, the screen needed to be shifted to 9 different locations in order to cover the whole camera grid. A total of 136 PSFs per channel were estimated for the Ximea camera. PSFs of all channels are overlaid and illustrated in Fig. 6. In a similar way, the process on the Blackberry phone's camera generated 117 PSFs shown in Fig. 1.

The measured PSFs along with sample images captured with these cameras were passed to a deconvolution algorithm. We applied Heide *et al.*'s deconvolution algorithm [10] as it handles chromatic artifacts successfully by employing a cross-channel prior. Fig. 7 shows the deconvolution results using the measured PSFs applied on the images captured by the Ximea and the Blackberry cameras. These results demonstrate how the measured lens PSFs are used to significantly enhance the quality of the images captured by the cameras.

Limitations Since lens PSF vary with depth, PSF estimation needs to be performed for different depths. In case of close-up PSF estimation, in order to avoid pixelation effects, a screen with high pixel density (PPI) is required. Moreover, to reduce the unwanted blur caused by the warping procedure, inverse mapping should be included in the warping function.

5. Conclusions

We proposed a new framework to estimate intrinsic camera lens blur. The proposed camera-scene alignment benefits from a high-resolution display to expose the calibration patterns. The fixed setup between the camera and the display allows us to switch different patterns and capture their images in a fixed geometric alignment. Hence, the calibration pattern can be separated from the pattern used in the
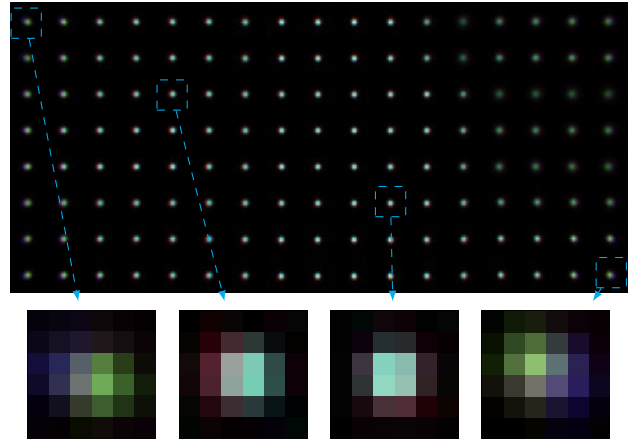

Figure 6. Lens PSFs measured for the Ximea camera.

PSF estimation. As a result, there is more flexibility to provide a large number of feature points in the calibration pattern and to guide the alignment more precisely. The warping procedure is reduced to a simple texture mapping due to appropriate number of feature points. Also, this fixed camera-scene alignment is used to produce intensity reference images to have pixel to pixel color correction in generating the sharp correspondence of the target image. Our PSF estimation method benefits from the frequency specifications of Bernoulli noise patterns to introduce a SDF constraint for the PSF. It is then used jointly with regularization terms in a non-negative constrained linear system to generate accurate lens PSFs. Experimental results show that our method is robust against noise, and therefore suitable for mobile devices. Our technique achieves better performance than the existing non-blind PSF estimation approaches.

Acknowledgment

References

[1] J. Brauers, C. Seiler, and T. Aach. Direct PSF estimation using a random noise target. In *IS&T/SPIE Electronic Imaging*, pages 75370B–75370B, 2010. 1, 2, 4

[2] D. C. Brown. Close-range camera calibration. *Photogramm. Eng.*, 37:855–866, 1971. 2

[3] T. Chan and C.-K. Wong. Total variation blind deconvolution. *IEEE Transactions on Image Processing*, 7(3):370–375, Mar 1998. 1

[4] S. Cho and S. Lee. Fast motion deblurring. *ACM Transactions on Graphics (SIGGRAPH)*, 28(5):145, 2009. 1

[5] T. S. Cho, S. Paris, B. K. Horn, and W. T. Freeman. Blur kernel estimation using the radon transform. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 241–248, 2011. 1

[6] M. Delbracio, A. Almansa, J.-M. Morel, and P. Musé. Sub-pixel point spread function estimation from two photographs
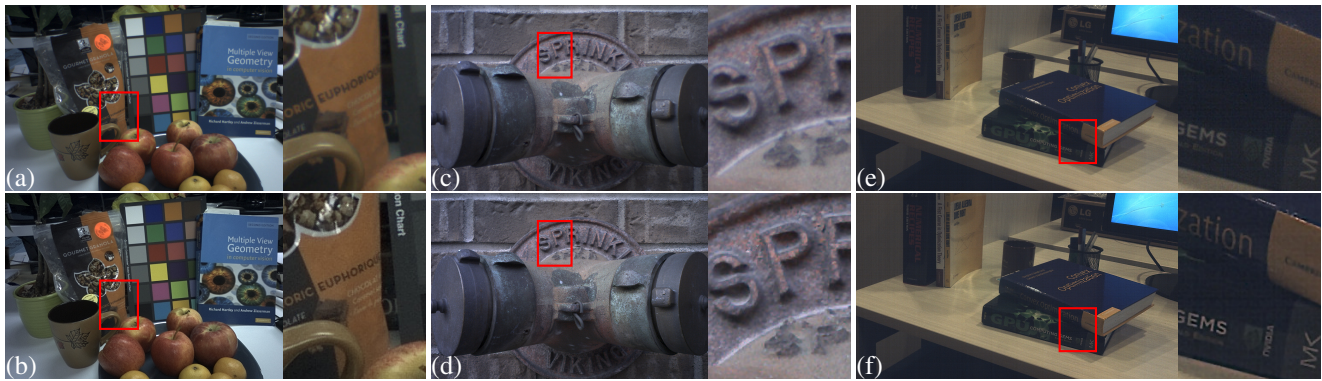
Figure 7. Debluring using estimated PSFs. (a,c) Images captured by the Blackberry phone's camera. (b,d) Deblurring using the measured lens PSFs shown in Fig. 1. (e) Image captured by the Ximea camera. (f) Deblurring using the measured lens PSFs shown in Fig. 6.

at different distances. *SIAM Journal on Imaging Sciences*, 5(4):1234–1260, 2012. 1, 2, 6

[7] M. Delbracio, P. Musé, A. Almansa, and J.-M. Morel. The non-parametric sub-pixel local point spread function estimation is a well posed problem. *International Journal of Computer Vision*, 96:175–194, 2012. 2, 4, 5, 6

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

[9] A. Goldstein and R. Fattal. Blur-kernel estimation from spectral irregularities. In *European Conference on Computer Vision (ECCV)*, pages 622–635, 2012. 1

[10] F. Heide, M. Rouf, M. B. Hullin, B. Labitzke, W. Heidrich, and A. Kolb. High-quality computational imaging through simple lenses. *ACM Transactions on Graphics (SIGGRAPH)*, 2013. 1, 2, 7

[11] N. Joshi, R. Szeliski, and D. Kriegman. PSF estimation using sharp edge prediction. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–8, 2008. 1, 2, 5, 6

[12] E. Kee, S. Paris, S. Chen, and J. Wang. Modeling and removing spatially-varying optical blur. In *IEEE International Conference on Computational Photography (ICCP)*, pages 1–8, 2011. 1, 2, 5, 6

[13] D. Krishnan, T. Tay, and R. Fergus. Blind deconvolution using a normalized sparsity measure. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 233–240, 2011. 1

[14] A. Levin. Blind motion deblurring using image statistics. In *Advances in Neural Information Processing Systems (NIPS)*, pages 841–848, 2006. 1

[15] A. Levin, P. Sand, T. S. Cho, F. Durand, and W. T. Freeman. Motion-invariant photography. In *ACM Transactions on Graphics (SIGGRAPH)*, pages 71:1–71:9, 2008. 1

[16] W. Li, J. Zhang, and Q. Dai. Exploring aligned complementary image pair for blind motion deblurring. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 273–280, June 2011. 1

[17] T. Michaeli and M. Irani. Blind deblurring using internal patch recurrence. In *European Conference on Computer Vision (ECCV)*, pages 783–798, 2014. 1

[18] Q. Shan, J. Jia, and A. Agarwala. High-quality motion deblurring from a single image. *ACM Transactions on Graphics (SIGGRAPGH)*, 27(3):73:1–73:10, Aug. 2008. 1

[19] Y. Shih, B. Guenter, and N. Joshi. Image enhancement using calibrated lens simulations. In *European Conference on Computer Vision (ECCV)*, pages 42–56. 2012. 1

[20] E. Simoncelli. Statistical models for images: compression, restoration and synthesis. In *Conference Record of the Thirty-First Asilomar Conference on Signals, Systems amp; Computers.*, volume 1, pages 673–678 vol.1, Nov 1997. 1

[21] J. Simpkins and R. L. Stevenson. Parameterized modeling of spatially varying optical blur. *Journal of Electronic Imaging*, 23(1):013005–013005, 2014. 1

[22] J. D. Simpkins and R. L. Stevenson. Robust grid registration for non-blind PSF estimation. In *Proc. SPIE Visual Information Processing and Communication*, volume 8305, pages 83050I–83050I–8, 2012. 2

[23] L. Sun, S. Cho, J. Wang, and J. Hays. Edge-based blur kernel estimation using patch priors. In *International Conference on Computational Camera (ICCP)*, 2013. 1

[24] M. Trimeche, D. Paliy, M. Vehvilainen, and V. Katkovnic. Multichannel image deblurring of raw color components. In *SPIE Computational Imaging*, pages 169–178, 2005. 1, 2

[25] L. Xu and J. Jia. Two-phase kernel estimation for robust motion deblurring. In *European Conference on Computer Vision (ECCV)*, pages 157–170. Springer, 2010. 1

[26] Y.-L. You and M. Kaveh. A regularization approach to joint blur identification and image restoration. *IEEE Transactions on Image Processing*, 5(3):416–428, Mar 1996. 1

[27] L. Yuan, J. Sun, L. Quan, and H.-Y. Shum. Image deblurring with blurred/noisy image pairs. *ACM Transactions on Graphics (SIGGRAPH)*, 26(3):1–10, 2007. 1

[28] T. Yue, S. Cho, J. Wang, and Q. Dai. Hybrid image deblurring by fusing edge and power spectrum information. In *European Conference on Computer Vision (ECCV)*, pages 79–93, 2014. 1

[29] J. Zandhuis, D. Pycock, S. Quigley, and P. Webb. Sub-pixel non-parametric PSF estimation for image enhancement. In *IEE Proceedings- Vision, Image and Signal Processing,*, volume 144, pages 285–292, 1997. 1