

From local to global: Edge profiles to camera motion in blurred images

Subeesh Vasu,¹ A.N. Rajagopalan²

Department of Electrical Engineering, Indian Institute of Technology Madras.

`subeeshvasu@gmail.com`¹, `raju@ee.iitm.ac.in`²

This is a draft version of the paper accepted for presentation at Conference on Computer Vision and Pattern Recognition (CVPR) JULY 2017. The final copyrighted version is available at

<https://ieeexplore.ieee.org/document/8099550/>

From local to global: Edge profiles to camera motion in blurred images

Subeesh Vasu¹, A. N. Rajagopalan²

Indian Institute of Technology Madras

subeeshvasu@gmail.com¹, raju@ee.iitm.ac.in²

Abstract

In this work, we investigate the relation between the edge profiles present in a motion blurred image and the underlying camera motion responsible for causing the motion blur. While related works on camera motion estimation (CME) rely on the strong assumption of space-invariant blur, we handle the challenging case of general camera motion. We first show how edge profiles ‘alone’ can be harnessed to perform direct CME from a single observation. While it is routine for conventional methods to jointly estimate the latent image too through alternating minimization, our above scheme is best-suited when such a pursuit is either impractical or inefficient. For applications that actually favor an alternating minimization strategy, the edge profiles can serve as a valuable cue. We incorporate a suitably derived constraint from edge profiles into an existing blind deblurring framework and demonstrate improved restoration performance. Experiments reveal that this approach yields state-of-the-art results for the blind deblurring problem.

1. Introduction

Camera motion estimation (CME) plays a vital role in many multi-image based applications such as structure from motion, super-resolution, HDR imaging etc. These applications involve sharing of information across images, which is typically achieved by establishing feature correspondences. But this becomes a challenging task in the presence of motion blur. CME is also important in many single motion blurred image based applications such as image restoration, scene depth estimation ([18, 36]), splicing detection ([12]), estimation of 3D structure of light sources ([38]) etc.

A common approach for CME from a motion blurred image is to pose it as a blind deblurring problem, wherein both latent image (or its features) and camera motion are estimated within an alternating minimization (AM) framework. In the blind deblurring literature, most approaches assume space invariant (SI) blur (i.e. pure in-plane camera translations) ([3, 9, 27, 10, 11, 1, 35, 15, 17, 2, 4, 29, 20, 21]). While this leads to very efficient blind deblurring algo-

rithms, this assumption does not hold true in practice. It is well-known ([16, 13]) that natural camera shake often contains rotational components resulting in spatially varying (SV) blur in the captured images. As is evident from recent works ([5, 31, 7, 34, 37, 22]), one needs to account for the SV nature of blur for accurate estimation of general camera motion. While this is an important problem, the increase in number of unknowns makes it quite ill-posed, demanding stronger priors for achieving quality performance. Very ill-posed problems such as depth-aware deblurring and rolling shutter motion deblurring, either avoid using AM altogether ([24]) or need a good initialization of the camera motion ([36, 28]) derived from local point spread functions (PSFs) i.e., blur kernel estimates returned by an off-the-shelf blind deblurring algorithm. However, their performance is limited because PSF estimates from small regions can be erroneous.

A recent trend in blind deblurring is to come up with class specific or generic priors ([35, 2, 29, 37, 20, 21, 22]) to reduce the ill-posedness of the problem. While some of these priors have shown great promise, their performance is limited due to one or more of the following reasons: failure on images for which prior does not hold good, inability to handle SV blur, computational complexity etc. In SV deblurring, [37] has been widely popular for a long time. They employed an L_0 prior on the gradients of the image while performing CME. The state-of-the-art work in [22] employs dark channel prior in addition to the L_0 prior of [37] to significantly improve the performance of SV deblurring.

In this work, we address the problem of CME from a single motion blurred image by harnessing edge profiles. There are only two notable works that have used edge profiles to derive useful information. Cho et al. [2] have modeled the relation between PSF and edge profiles to perform SI deblurring. Tai et al. [30] have used edge profiles for estimation of camera response function (CRF) from SI blurred images. We first propose an approach (called **EpAlone**) for efficient and direct CME from edge profiles i.e., sans the need for latent image estimation. We expand the scope of edge profiles and employ them for estimation of general camera motion. While problems such as blind deblurring,

HDR imaging etc. perform joint estimation of the latent image and camera motion, for applications where such a scheme is impractical or inefficacious, **EpAlone** is apt. For applications that actually rely on such a joint estimation scheme, a wise choice is to use edge profiles as an additional regularizer. Recent works on blind deblurring have ([21, 22]) revealed that the performance can be improved by adding strong priors onto existing methods. Following these works, we propose a method (called **Xu + Ep**) that incorporates an efficient and effective constraint derived from edge profiles into the existing method of Xu et al. [37] to achieve state-of-the-art performance. Both the proposed methods (**EpAlone** and **Xu + Ep**) are based on new theoretical claims that we introduce in this paper. The corresponding proofs are provided in the supplementary material due to space constraints.

Our main contributions are summarized below.

- For the first time in literature, we derive and analyze the relation between edge profiles present in a blurred image and the underlying 6D camera motion.
- We propose a new constraint derived from edge profiles and use it to perform efficient and direct CME from a single motion blurred image.
- We propose a blind deblurring scheme by elegantly incorporating our new constraint into the formulation of [37] to achieve state-of-the-art performance.

2. Edge profile and convolution

Edge profile is the alpha matte taken over a line along the edge orientation (by treating the two colors of homogeneous regions on either sides of the edge as background and foreground). In our work, the boundary between two homogeneous regions in an image is considered as an edge. Let us denote the edge profile of length N across an edge as α_i where $i = 1, 2, \dots, N$. Then the intensity values along a line L lying along the edge orientation can be expressed as

$$L(i) = c_1 \alpha_i + c_2 (1 - \alpha_i) \quad (1)$$

with c_1 and c_2 being the colors of the two homogeneous regions on either side of the edge. Without loss of generality, assume that the edge that we are working with is a vertical edge. Consider a binary step edge S defined over a rectangular region around the considered edge as

$$S(x, y) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0. \end{cases} \quad (2)$$

For a clean image, the intensity I at this region can be expressed as,

$$I = S c_1 + (1 - S) c_2 \quad (3)$$

Blurring the above region with a kernel k results in

$$B_I = k * I = (k * S) c_1 + (1 - k * S) c_2 \quad (4)$$

$$= S_k c_1 + (1 - S_k) c_2 \quad (5)$$

where S_k embeds the variation in intensity while moving across the edge in the motion blurred image B_I . By comparing Eq. (1), Eq. (2), and Eq. (5) we can observe that, the edge profile α in the blurred image is equal to the middle row of S_k . It is trivial to see that the rows of S_k and the cumulative of the projection of k onto X axis are equivalent, i.e.,

$$S_k(x, y) = \sum_{u=-\infty}^x \sum_{v=-\infty}^{\infty} k(u, v) = \sum_{u=-\infty}^x P(u) \quad (6)$$

where P is the projection of k onto the X axis. Note that the final expression in Eq. (6) is not a function of y indicating that all rows of S_k will be identical.

Similarly one can show that, in general, an edge profile is equivalent to the cumulative distribution function (CDF) of the projection of PSF onto the line along the edge orientation [30] (or CDF of the Radon transform of the PSF along the direction orthogonal to the edge orientation [2]). If the local region I in the image is not blurred (i.e., $B_I = I$), then $S_k = S$ i.e., the edge profile in that region will be a step response. Note that these properties of step edges are defined over local regions. Since we can treat the blur at local regions as SI even for general camera motion, we can use the above results in SV blurred images too. This opens up the possibility of exploiting edge profiles for general CME. More specifically, one could solve for the camera motion by looking for the solution which favors ideal step edges in corresponding latent image estimates.

An illustrative example supporting the results discussed in this section is displayed in Fig. 1. We have used a sample PSF (Fig. 1(a)) from the dataset of [16] to blur an example image, a region extracted from which is shown in Fig. 1(b). The projection of PSF onto two different directions along with the CDF of this projection is displayed in Fig. 1(a). As is evident from Fig. 1(b), the projection of PSF is equal to the differential of the edge profile extracted along the respective directions.

3. Edge profile to camera motion

In this section, we derive the relation between the local edge profiles present in a motion blurred image and the underlying global camera motion. For general camera motion, the blurred image (g) can be expressed as an aggregate of warped instances of the latent image f as ([5, 34, 23])

$$g(\mathbf{x}) = \frac{1}{t_e} \int_0^{t_e} f((H^t)^{-1}(\mathbf{x})) dt \quad (7)$$

where $\mathbf{x} = (x, y)$ represents the spatial coordinates, t_e is the total exposure time of the image, H^t corresponds to the homography relating the latent image to the projection of

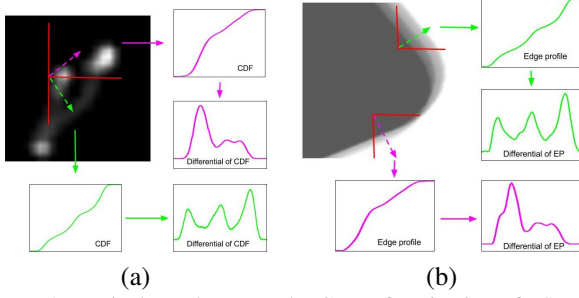


Figure 1. Equivalence between the CDF of projection of PSF and its corresponding edge profile formed in the blurred image. (a) A sample PSF and its projection along two directions and the CDF of the projections. (b) An example region from a corresponding blurred image along with the edge profile extracted from it (along orthogonal directions) as well as the differential of the edge profile.

the scene onto the image plane at the time instant t . An equivalent way to relate g and f is to use the warp which is defined based on the poses that the camera undergoes during motion.

$$g(\mathbf{x}) = \int_{\gamma \in T} \Psi(\gamma) f((H^\gamma)^{-1}(\mathbf{x})) d\gamma \quad (8)$$

where T is the set of different poses that the camera undergoes during the exposure time, γ refers to a single pose from T , H^γ is the homography warp corresponding to γ and $\Psi(\gamma)$ is the fraction of time over which the camera stays at γ . Eq. (8) allow to express the same relation as Eq. (7) but in a time-independent fashion. Ψ is referred to as motion density function (MDF) ([5]) or transformation spread function (TSF) ([23]). For general camera motion, the PSF at \mathbf{x} in the blurred image can be related to the MDF ([23]) as

$$k(\mathbf{x}, \mathbf{u}) = \int_{\gamma \in T} \Psi(\gamma) \delta(\mathbf{u} - (H^\gamma \mathbf{x} - \mathbf{x})) d\gamma \quad (9)$$

where $H^\gamma \mathbf{x}$ refers to $H^\gamma(\mathbf{x})$, the coordinates obtained by warping \mathbf{x} with the homography H^γ . Here PSF $k(\mathbf{x}, \mathbf{u})$ is defined as a function of two variables, since it is space variant (SV) in nature, where $\mathbf{u} = (u, v)$ represents the spatial coordinates over which the PSF at \mathbf{x} is defined. Eq. 9 is a direct implication of the fact that a pose vector γ_0 assign its weight $\Psi(\gamma_0)$ to a point \mathbf{u}_0 in the PSF, and the PSF can be obtained by integrating such assignments from all poses in MDF. The pose γ_0 can be related to \mathbf{u}_0 as

$$\mathbf{u}_0 = H^{\gamma_0} \mathbf{x} - \mathbf{x} \quad (10)$$

Let us denote $k(\mathbf{x}, \mathbf{u})$ as $k_{\mathbf{x}}(u, v)$. The differential of edge profile $dE_{\theta, \mathbf{x}}$ (along $\theta + \pi/2$) at \mathbf{x} can be expressed as the Radon transform of the PSF at \mathbf{x} along θ ([2]).

$$dE_{\theta, \mathbf{x}}(\rho) = \int \int k_{\mathbf{x}}(u, v) \delta(\rho - u \cos \theta - v \sin \theta) du dv \quad (11)$$

where ρ represents the radial coordinates of the Radon projection. Combining the relations in Eq. (9) and Eq. (11) we are able to directly relate MDF to the edge profile as¹

$$dE_{\theta, \mathbf{x}}(\rho) = \int_{\gamma \in \Gamma} \Psi(\gamma) \delta(\rho - (H_x^\gamma \mathbf{x} - x) \cos \theta - (H_y^\gamma \mathbf{x} - y) \sin \theta) d\gamma \quad (12)$$

Here $H_x^\gamma \mathbf{x}$ refers to the x coordinate of $H^\gamma(\mathbf{x})$. Eq. (12) relates the differential of edge profile (dE) directly with the MDF or the underlying camera motion. By discretizing the relation in Eq. (12) we obtain

$$dE_{\theta, \mathbf{x}}(\rho) = \sum_{p=1}^{N_T} w(p) \delta(\rho - (H_x^p \mathbf{x} - x) \cos \theta - (H_y^p \mathbf{x} - y) \sin \theta) \quad (13)$$

where w is the MDF weight vector defined over a discrete pose space P with the value being nonzero (and positive) only for those poses $p \in P$ over which the camera has moved, and N_T is the total number of poses in P . This discretization allows us to relate the edge profile with the MDF vector w in matrix-vector multiplication form as

$$dE_{\theta, \mathbf{x}} = M_{\theta, \mathbf{x}} w \quad (14)$$

where $M_{\theta, \mathbf{x}}$ is the measurement matrix formed according to the relation in Eq. (13), $dE_{\theta, \mathbf{x}} \in N_\rho \times 1$, $M_{\theta, \mathbf{x}} \in N_\rho \times N_T$ and $w \in N_T \times 1$ with N_ρ being the length of the edge profile. Using such N_e edge profiles, we can write

$$dE = M w \quad (15)$$

where $dE \in N_\rho N_e \times 1$, $M \in N_\rho N_e \times N_T$. Thus a collection of edge profiles from a motion blurred image can be related to the MDF elegantly through Eq. (15). We can solve for MDF w by minimizing the following cost function.

$$\hat{w} = \arg \min_w \|dE - M w\|_2 + \lambda \|w\|_1 \quad (16)$$

where L_1 norm is used to enforce the natural sparsity of camera motion trajectory in the pose space, and λ is the prior weight on MDF. Note that this approach is similar to the idea of recovering MDF from local PSF estimates as suggested in [24]. Since the local PSF estimates obtained through conventional blind deblurring schemes can be erroneous, the applicability of CME from PSF is limited. On the other hand, our formulation in Eq. (16) directly relate camera motion to the naturally available edge profiles in the image, improving the scope of direct CME. In other words, while the performance of CME from PSFs depends on the performance of existing methods on PSF estimation, our method does not suffer from such a limitation.

¹Details of the derivation can be found in supplementary.

4. Recovering camera motion from edges

If sufficient number of edge profiles is available, we can use Eq. (16) to estimate camera motion accurately. However, there are some issues in attempting to use Eq. (16) in practical situations. To extract edge profiles from different locations in an image, we follow the procedure in [2]. They have used a set of image analysis heuristics including detection of edges in blurry images and a number of outlier rejection schemes. However, a fundamental problem is that the extracted edge profiles will not be aligned with respect to each other, and hence cannot be directly used for CME in Eq. (16). To align the extracted edge profiles, [2] exploits the fact that the center of mass of the PSF corresponds to the center of mass of their 1D projections. Since this was not valid under the influence of CRF, [30] employed extreme points to align the edge profiles along the same direction. But this is not valid for edge profiles from different orientations, since the homographies which map to extreme points can be different for projections along different directions.² Hence, for our scenario we cannot use extreme points for alignment.

4.1. Alignment and implications

In this section, we show that edge profiles can be aligned using the center of mass, even for general camera motion. We restrict our analysis to two different but commonly used approximations for 6D camera trajectories. First, we analyze the case of *model*₁ which models 6D motion using in-plane translations and in-plane rotations ([5]). Later we extend the result to *model*₂ which employs pure 3D rotations ([34]) as an approximation to 6D motion.

Claim 1: For *model*₁, if the in-plane rotational motion undergone by the camera is small (as is typical for camera shakes in handheld situations [13, 28]) then it can be shown that the centroid of MDF vector will correspond to the centroid of PSF i.e., the centroid pose of the camera obtained from MDF will map points to the centroid of PSFs generated by the camera motion or equivalently to the centroid of the corresponding edge profiles.

For small in-plane rotations one can show that the centroid co-ordinates (u_c, v_c) of PSF at \mathbf{x} in the image corresponds to the centroid $(t_x^c, t_y^c, \theta_z^c)$ of MDF.³ i.e.,

$$u_c = \sum_u \sum_v u k_{\mathbf{x}}(u, v) = H_x^c \mathbf{x} - x \quad (17)$$

$$v_c = \sum_u \sum_v v k_{\mathbf{x}}(u, v) = H_y^c \mathbf{x} - y \quad (18)$$

where H^c is the homography corresponding to the centroid of MDF. The correspondence between the centroid of a PSF

²An illustrative example is provided in supplementary.

³Detailed proof is provided in supplementary.

and its corresponding edge profiles is trivial to see from the fact that the edge profile is a linear projection of PSF. Since the correspondence between the centroids of edge profile and PSF is always valid [2], the alignment of PSFs automatically refers to alignment of edge profiles.

Furthermore, since in-plane translations are equivalent to out-of-plane rotations (for large focal length) [34], we can conclude that the centroid based alignment of edge profiles is sufficient for *model*₂ also. This is also supported by the work in [13] which shows that *model*₁ and *model*₂ are equally good in approximating the original 6D camera trajectories caused by natural handshakes.

Next, we analyze the effect of centroid-based alignment on the estimation of camera motion and latent image. It is well-known that the same blurred image can be produced by different valid (i.e., correct upto a homography) latent image-camera motion pairs. The aligned edge profiles should correspond to a valid camera motion to ensure that the estimate returned by Eq. (16) is correct. Next, we will show that a collection of centroid aligned PSFs is consistent with a valid camera motion and latent image that can produce the same blurred image (implying that the aligned edge profiles can be directly used for CME). Let us consider a collection of m PSFs, k_1, \dots, k_m obtained from locations $\mathbf{x}_1, \dots, \mathbf{x}_m$ in the blurred image.

Claim 2: The centroid aligned PSFs correspond to a camera motion and a latent image both of which are consistent with the given blurred image and correct upto a homography.

Let there be n active (non-zero) poses in the MDF w_0 , with H^1, \dots, H^n being the corresponding homographies, and L be the latent image. We can show that⁴ the aligned PSFs corresponds to another latent image (L^c) defined on coordinates $H^c \mathbf{x}$ (i.e. related to the L through the homography warp H^c) and a new MDF (formed of a collection of homographies $H^j (H^c)^{-1}$) both of which are consistent with the same blurred image. Thus, using the centroid aligned edge profiles we can estimate one of the valid camera motion w_0^c using Eq. (16).

To verify the results of claims 1 and 2, and also to analyze the impact of the error introduced in the CME by centroid alignment of edge profiles, we performed the following experiment. Using camera trajectories (ground truth camera motion) from the database of [13], we produced edge profiles at different locations in the image. These edge profiles are aligned with respect to the centroid and then used for estimation of corresponding MDF using Eq. (16). This MDF is then used to generate PSFs at different locations. We generate another set of PSFs using the centroid aligned camera trajectories obtained by mapping the ground truth camera trajectory using $(H^c)^{-1}$. From claims 1 and 2, these two PSFs should be identical. We computed the

⁴For detailed discussions on claim 2, refer to supplementary material.

normalized cross correlation (NCC) values between these two sets of PSFs to verify our observations. This also measures the accuracy (validity) of centroid alignment of PSFs for different kinds of camera motions. In addition to this, we computed the difference (say d_c) between the centroid of the edge profiles and the point to which the centroid of camera trajectories maps. We carry out experiments in two different ways. On the one hand, we tested the applicability of our proposed method on $model_1$ and $model_2$ using the camera trajectories from database of [13] (Figs. 2(b, c, e, f)). On the other hand, we separately investigated the variation of the accuracy with respect to in-plane rotations alone using a single camera trajectory. The amount of rotation spanned by the trajectory is increased to observe the behavior of both error in centroid as well as NCC value (displayed in Figs. 2(a, d)). We found that d_c is always sub-pixel (Figs. 2(a,b,c)) and the correlation values are consistently high (Figs. 2(d,e,f)), which supports our theoretical findings. Although the error increases with increase in the degree of rotation, it is still quite small even for in-plane rotations upto 7 degrees, which is well within the range introduced by incidental camera shakes.

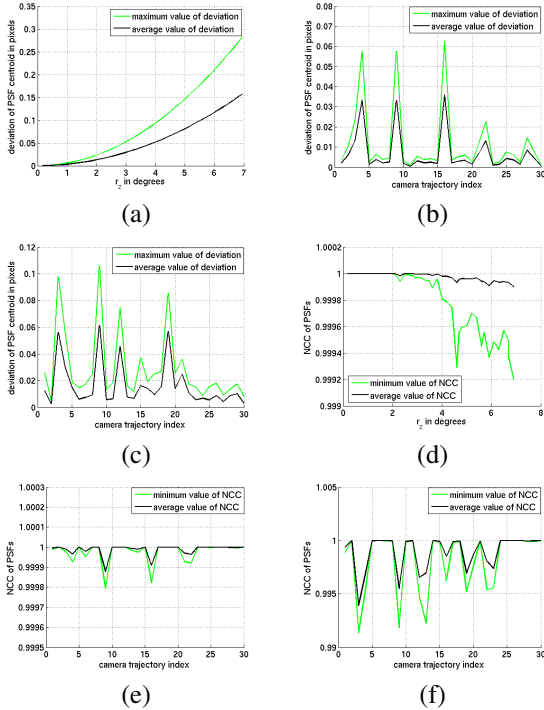


Figure 2. Analyzing the impact of centroid alignment on CME. These measures are estimated from PSFs generated all over an example image. Absolute value of maximum and average error between estimated centroid and ground truth centroid for the camera motion using (a) in-plane rotation only, (b) $model_1$, and (c) $model_2$. Minimum and average value of NCC between non-blindly estimated PSFs and GT PSFs for (d) in-plane rotation only, (e) $model_1$, and (f) $model_2$.

4.2. Computational considerations

On the basis of claims 1 and 2, we can align all the edge profiles extracted from a given motion blurred image and perform CME. However, direct use of Eq. (16) is computationally expensive. This is because both the construction of the measurement matrix M and operations involving such a large matrix are time-consuming. Therefore, we propose to re-frame our constraint in Eq. (16) so as to exploit the advantages of the computationally efficient filter-flow framework ([8]).

From claims 1 and 2, we observe that the centroid locations of edge profiles on the blurred image can be treated as the location of the step edges in the corresponding latent image. From this principle, we can generate a prediction image (I_p) formed from all the step edges in the latent image. We build an edge profile image (I_{ep}) formed from the differential of edge profiles at the respective locations. Using Eq. (6), Eq. (15) and claims 1 and 2, we note that the step edges in I_p are related to the corresponding edge profiles in I_{ep} through a valid camera motion. Hence, we propose to use the following estimation equation, formed from I_p and I_{ep} , to replace the computationally expensive formulation of Eq. (16).

$$\hat{w} = \arg \min_w ||T_w(\tilde{I}_p) - \tilde{I}_{ep}||_2 + \lambda ||w||_1 \quad (19)$$

Here T_w represents the set of warping operations defined by w . This non-uniform blur relation between I_p and I_{ep} can be implemented efficiently using the filter-flow framework [8]. We use our proposed constraint in Eq. (19) to perform blind CME from edge profiles and name this as **EpAlone**. The computational complexity of **EpAlone** can be reduced further, as discussed next in the generation of I_p and I_{ep} .

Claim 3: The absolute value of differential of edge profile (P in Eq. (6)) is equivalent to the absolute gradient of a blurred image (at the location of the edge profile) normalized by the difference between the two homogeneous colors involved. i.e.,

$$|\nabla B_I| = |(c_1 - c_2)|P \quad (20)$$

From claim 3,⁵ we form I_p by assigning the absolute value of $(c_1 - c_2)$ at the centroid positions. I_{ep} is formed by collecting the slices (around the centroid of edge profile) of absolute gradients along the direction of edge profiles. In our experiments, we observed that taking slices only along the direction of edge profile often leads to small errors in the estimate of camera motion, mainly due to errors in the estimate of edge profile orientation. As a solution, we have chosen slices over a small angular window around the estimated edge profile orientation. This improves the robustness of Eq. (19) by making it less sensitive to errors in orientation estimates (as compared to Eq. (16)).

⁵For detailed discussions on claim 3, please refer to supplementary.

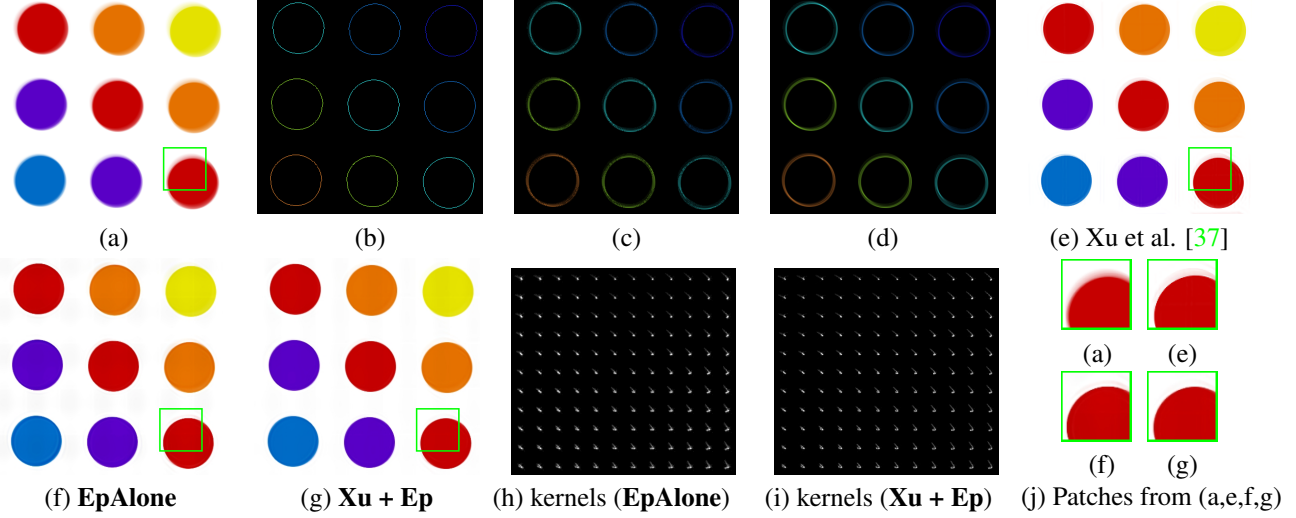


Figure 3. Synthetic example on non-uniform deblurring: (a) Input motion blurred image, and (b) our prediction image. Edge profile image generated by (c) single-direction mask, and (d) multi-direction mask. (e-g) Deblurred results. (h-i) Estimated kernels from proposed methods. (j) Patches from (a,e,f,g)

An illustrative example on the generation of I_p and I_{ep} is given in Fig. 3. We have used the centroid locations of extracted edge profiles from the blurred image (Fig. 3(a)) to form the prediction image (Fig. 3(b)), with the value at each position being the absolute value of differences between the colors of nearby homogeneous regions. The corresponding edge profile images formed by using single as well as multi-direction mask is shown in Figs. 3(c-d).

Remarks: Our idea of building the prediction image is similar to the use of edge prediction units in existing deblurring approaches. The difference lies in the fact that, we use the idea of edge profiles for direct prediction of step edges, as compared to the iterative prediction schemes (through filtering techniques or image priors) used by other methods. Our primary objective in proposing **EpAlone** is to reveal the potential of the new constraint (that we advocate through Eq. (19)) ‘alone’ to regularize the problem of CME.

5. Deblurring with edge profile constraint

Since the dimensionality of unknowns is large, Eq. (19) needs adequate number of edge profiles spanning a wide range of orientations to obtain an accurate estimate of MDF. But it may not always be possible to fulfill this requirement. In this section, we show that Eq. (19) can still serve as a valuable constraint to regularize existing methods for CME. This is in spirit with recent works on blind deblurring ([21, 22]) that have proposed new priors as an addition to the image prior in [37], to achieve better performance. Following this line of thought, we also propose to use Eq. (19) as an additional constraint to the existing method in [37] to uplift its performance. Although our constraint can be used to improve the performance of any deblurring meth-

ods, we chose to use [37], since their L_0 norm prior on latent image gradients was found to yield the best results among works in SV deblurring. We incorporate Eq. (19) into the formulation of [37] and call this new approach as **Xu + Ep**. Following are the modified equations used for estimation of latent image and MDF in **Xu + Ep**.

$$\tilde{f}^{t+1} = \arg \min_{\tilde{f}} \{ \|M_w^t \tilde{f} - \tilde{g}\|_2 + \lambda_f \|\nabla \tilde{f}\|_0 \} \quad (21)$$

$$w^{t+1} = \arg \min_w \{ \|M_{\nabla f}^{t+1} w - \nabla \tilde{g}\|_2 + \lambda_{I_p} \|M_{I_p} w - \tilde{I}_{ep}\|_2 + \lambda_w \|w\|_1 \} \quad (22)$$

where t is the iteration number, $\tilde{*}$ refers to lexicographically arranged form of $*$. $M_{\nabla f}$ and M_{I_p} are the warping matrices built from gradient of latent image (∇f) and I_p , and M_w is the measurement matrix formed of w . The first and second term in Eq. 22 act as data constraints whereas the last term is the sparsity prior on the camera trajectory. While λ_{I_p} controls the influence of our proposed constraint on the entire optimization, λ_w determines the degree of sparsity in w . In Eq. 21, we have used an L_0 norm prior ([37, 21, 22]) on image gradients weighted by a scalar parameter λ_f for latent image estimation. It is straightforward to see that by enforcing our edge-profile constraint, we are indirectly motivating the step edges present in the latent image to strongly influence the camera motion estimate. This improvement over [37] delivers state-of-the-art performance as we shall show later.

Since our additional constraint respects the filter-flow framework formulation ([7, 32]), the optimization problems in Eq. (21) and Eq. (22) can be solved entirely using filter-flow. In **Xu + Ep**, we alternately minimize between latent image and MDF estimation in a scale-space fashion. All

Algorithm 1 CME for a single scale.

Input: Blurred image g at current scale, initial MDF estimate w^0 obtained from previous scale.

Output: Refined MDF estimate at current scale.

- 1: **For** $t = 1 : n_{max}$
 - 2: Estimate f^t by solving Eq. (21)
 - 3: Estimate w^t by solving Eq. (22)
 - 4: **End**
 - 5: Return w^t as the refined estimate of MDF for current scale.
-

major steps (for a single scale) for **Xu + Ep** is listed in Algorithm 1.⁶ It should be noted that, the objective of the AM is to obtain an accurate estimate of camera motion. To obtain the final restored image, we perform a non-blind deblurring using the camera motion estimate returned by AM.

Remarks: Exploiting claims 1, 2, and 3 we have developed a computationally efficient constraint leading to two elegant CME methods, **EpAlone** and **Xu + Ep**. However, there are important differences in their applicability. Unlike existing approaches (including **Xu + Ep**), **EpAlone** circumvents the need for intermediate latent image estimation. Therefore, applications such as CRF estimation ([30]), normal recovery ([25]), depth-aware deblurring ([24, 36]), rolling shutter motion deblurring [28] etc. that either do not warrant AM or need a good initial estimate of camera motion, can benefit from **EpAlone** or its underlying formulations. For applications such as blind deblurring, HDR imaging etc. which rely on AM and explicit latent image estimation, the edge profile constraint we derived can be a valuable add-on as discussed for the deblurring problem.

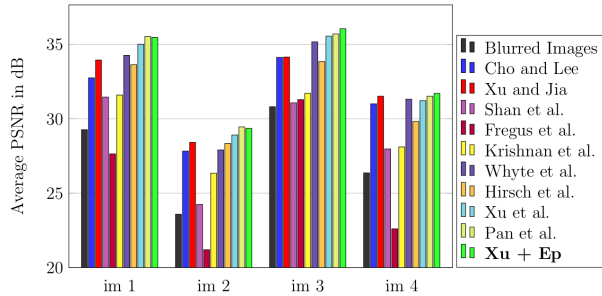


Figure 4. Quantitative evaluation on first 7 kernels from [13].

6. Experiments

Since visual evaluation of estimated camera motion is difficult, we compare the performance of our proposed methods based on the latent image estimates obtained from the resulting camera motion. This in turn signifies the ac-

⁶Implementation details of our algorithm and the parameter settings used in our experiments are provided in the supplementary.

curacy of CME. We use [14] to get a latent image estimate using the MDF obtained from both **EpAlone** and **Xu + Ep**. We employ the filter-flow implementation of Lasso algorithm [19] provided by [33, 32], for CME using both **EpAlone** and **Xu + Ep**. Although our main focus was on CME from non-uniformly blurred images, we do comparisons for SI cases too. For SI blur, we limit our qualitative comparisons to two closely related approaches [2, 37], and the current state-of-the-art work [22]. For SV blur, we compare with [5, 34, 6, 7, 37, 22], and [26]. Qualitative comparisons are done mainly on real image databases from [2, 5, 34, 6], and [22]. For fairness, quantitative comparisons are done only for AM-based deblurring methods.

Quantitative evaluation: We have used 28 blurred images formed from the 4 latent images and first 7 kernels in the benchmark dataset in [7]. We excluded other kernels since they represent very large blurs which can lead to failure of edge profile extraction. As is evident from the quantitative comparisons in Figure 4, **Xu + Ep** outperforms competing methods ([1, 35, 27, 3, 15, 32, 7, 37]) and is comparable or even better than state-of-the-art ([22]).

Synthetic example: Figure 3 shows a synthetic example on non-uniform deblurring. This is an ideal example for edge profile based deblurring, because of abundant availability of edge profiles spanning all orientations. This enables **EpAlone** Fig. 3(d) to perform better than the competing methods. Fig. 3(e) is the result from **Xu + Ep**, which gives the best result of all.

Real examples: We give few representative examples on deblurring here (Fig. 5). More examples are provided in the supplementary. For completeness and to demonstrate the potential of **EpAlone**, we have included the results from **EpAlone** for the application of blind deblurring. Note that in most of the examples, **EpAlone** by itself deliver good quality output, although it uses only edge profiles for CME. Our proposed approach (**Xu + Ep**) is able to give results with comparable or better quality to the state-of-the-art for all examples. A visual comparison also reveals that **Xu + Ep** is able to integrate the strength of [37] and **EpAlone** to achieve the best performance. For the examples in Fig. 3 and Fig. 5, [37] (Fig. 3(e), Fig. 5(b)) leaves residual blur in the estimated latent image, whereas **EpAlone** (Fig. 3(f), Fig. 5(d)) produces ringing artifacts while trying to enforce abrupt transitions across step edges. However, **Xu + Ep** (Fig. 3(g), Fig. 5(e)) is able to integrate desirable features from both methods to capture the motion accurately.

Remarks: **EpAlone** performs CME faster than both **Xu + Ep** and [22], while producing promising results. At the same time our efficient and improvised form of edge constraint enables **EpAlone** to perform better than its SI counterpart in [2]. The main advantage of **Xu + Ep** over the state-of-the-art work ([22]) is the significant reduction in computational complexity, but with comparable/improved

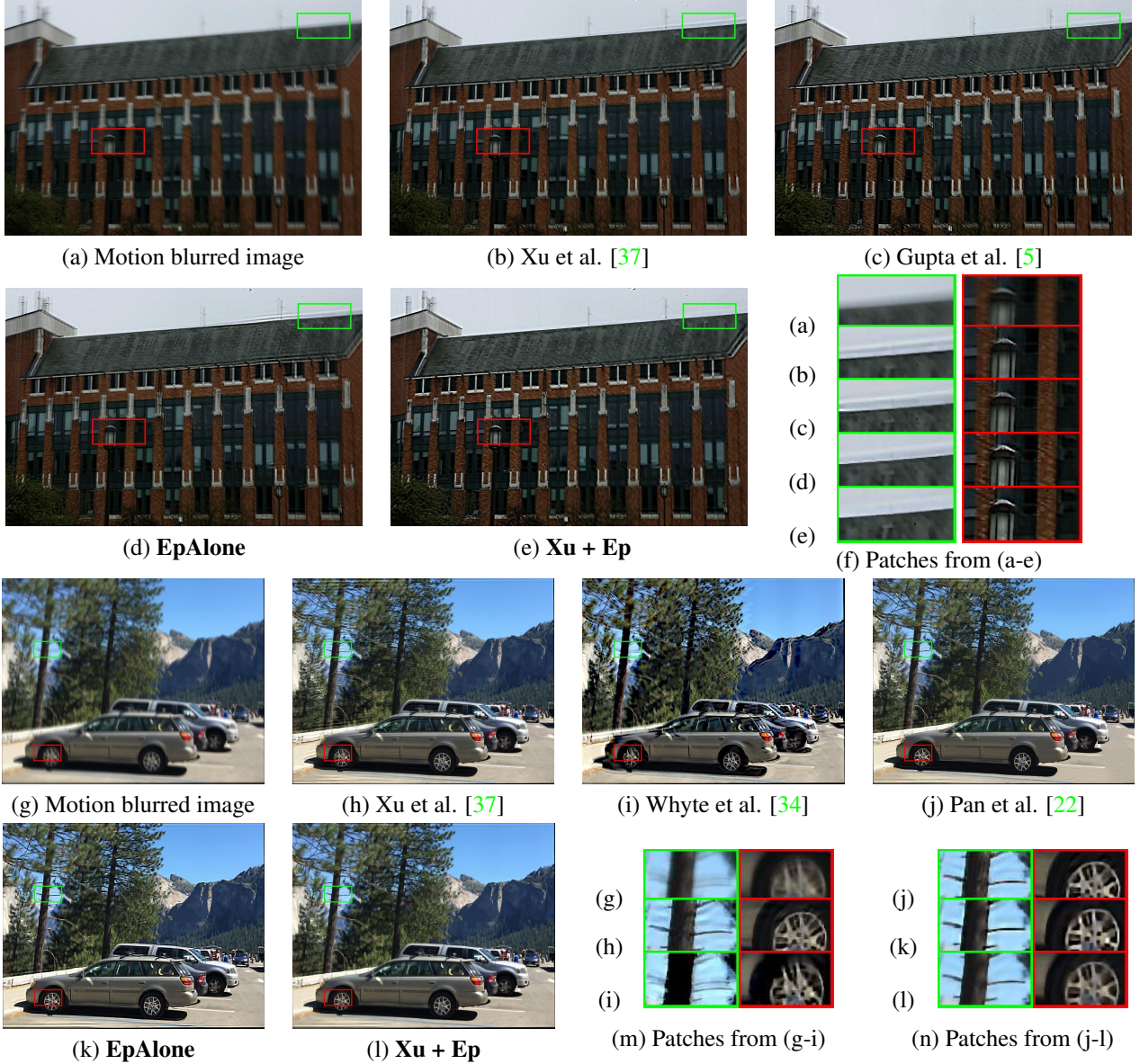


Figure 5. Real example on SV deblurring using image from the datasets of Gupta et al. [5] and Pan et al. [22].

performance. Comparisons of running times are provided in the supplementary. The main limitation of our new constraint is its inability to handle large blurs unlike [22].

7. Conclusions

In this paper, we investigated the relation between edge profiles present in a motion blurred image and the underlying camera motion. We proposed a method (**EpAlone**) for direct CME from edge profiles to aid applications where typical AM frameworks are inapt. To highlight the importance of our new edge constraint for applications that actually favor AM, we incorporated it into an existing blind deblurring framework and demonstrated improved perfor-

mance. Experiments reveal that our proposed approach (**Xu + Ep**) yields state-of-the-art results for the blind deblurring problem, with significant improvements in speed-up. As future work, it will be interesting to use the results developed in our paper (directly or indirectly) for other important applications such as CRF estimation, normal estimation, depth-aware deblurring, rolling shutter motion deblurring etc.

References

- [1] S. Cho and S. Lee. Fast motion deblurring. In *ACM Transactions on Graphics (TOG)*, volume 28, page 145. ACM, 2009. 1, 7

- [2] T. S. Cho, S. Paris, B. K. Horn, and W. T. Freeman. Blur kernel estimation using the radon transform. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 241–248. IEEE, 2011. 1, 2, 3, 4, 7
- [3] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman. Removing camera shake from a single photograph. In *ACM Transactions on Graphics (TOG)*, volume 25, pages 787–794. ACM, 2006. 1, 7
- [4] A. Goldstein and R. Fattal. Blur-kernel estimation from spectral irregularities. In *European Conference on Computer Vision*, pages 622–635. Springer, 2012. 1
- [5] A. Gupta, N. Joshi, C. L. Zitnick, M. Cohen, and B. Curless. Single image deblurring using motion density functions. In *European Conference on Computer Vision*, pages 171–184. Springer, 2010. 1, 2, 3, 4, 7, 8
- [6] S. Harmeling, H. Michael, and B. Schölkopf. Space-variant single-image blind deconvolution for removing camera shake. In *Advances in Neural Information Processing Systems*, pages 829–837, 2010. 7
- [7] M. Hirsch, C. J. Schuler, S. Harmeling, and B. Schölkopf. Fast removal of non-uniform camera shake. In *2011 International Conference on Computer Vision*, pages 463–470. IEEE, 2011. 1, 6, 7
- [8] M. Hirsch, S. Sra, B. Schölkopf, and S. Harmeling. Efficient filter flow for space-variant multiframe blind deconvolution. In *CVPR*, volume 1, page 2, 2010. 5
- [9] J. Jia. Single image motion deblurring using transparency. In *2007 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8. IEEE, 2007. 1
- [10] N. Joshi, R. Szeliski, and D. J. Kriegman. Psf estimation using sharp edge prediction. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8. IEEE, 2008. 1
- [11] N. Joshi, C. L. Zitnick, R. Szeliski, and D. J. Kriegman. Image deblurring and denoising using color priors. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 1550–1557. IEEE, 2009. 1
- [12] P. Kakar, N. Sudha, and W. Ser. Exposing digital image forgeries by detecting discrepancies in motion blur. *IEEE Transactions on Multimedia*, 13(3):443–452, 2011. 1
- [13] R. Köhler, M. Hirsch, B. Mohler, B. Schölkopf, and S. Harmeling. Recording and playback of camera shake: Benchmarking blind deconvolution with a real-world database. In *European Conference on Computer Vision*, pages 27–40. Springer, 2012. 1, 4, 5, 7
- [14] D. Krishnan and R. Fergus. Fast image deconvolution using hyper-laplacian priors. In *Advances in Neural Information Processing Systems*, pages 1033–1041, 2009. 7
- [15] D. Krishnan, T. Tay, and R. Fergus. Blind deconvolution using a normalized sparsity measure. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 233–240. IEEE, 2011. 1, 7
- [16] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman. Understanding and evaluating blind deconvolution algorithms. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 1964–1971. IEEE, 2009. 1, 2
- [17] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman. Efficient marginal likelihood optimization in blind deconvolution. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 2657–2664. IEEE, 2011. 1
- [18] H.-Y. Lin and C.-H. Chang. Depth recovery from motion blurred images. In *18th International Conference on Pattern Recognition (ICPR'06)*, volume 1, pages 135–138. IEEE, 2006. 1
- [19] J. Liu, S. Ji, J. Ye, et al. Slep: Sparse learning with efficient projections. *Arizona State University*, 6:491, 2009. 7
- [20] T. Michaeli and M. Irani. Blind deblurring using internal patch recurrence. In *European Conference on Computer Vision*, pages 783–798. Springer, 2014. 1
- [21] J. Pan, Z. Hu, Z. Su, and M.-H. Yang. Deblurring text images via l0-regularized intensity and gradient prior. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2901–2908, 2014. 1, 2, 6
- [22] J. Pan, D. Sun, H. Pfister, and M.-H. Yang. Blind image deblurring using dark channel prior. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016. 1, 2, 6, 7, 8
- [23] C. Paramanand and A. Rajagopalan. Shape from sharp and motion-blurred image pair. *International journal of computer vision*, 107(3):272–292, 2014. 2, 3
- [24] C. Paramanand and A. N. Rajagopalan. Non-uniform motion deblurring for bilayer scenes. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1115–1122, 2013. 1, 3, 7
- [25] M. P. Rao, A. Rajagopalan, and G. Seetharaman. Inferring plane orientation from a single motion blurred image. In *Pattern Recognition (ICPR), 2014 22nd International Conference on*, pages 2089–2094. IEEE, 2014. 7
- [26] C. J. Schuler, M. Hirsch, S. Harmeling, and B. Schölkopf. Learning to deblur. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(7):1439–1451, 2016. 7
- [27] Q. Shan, J. Jia, and A. Agarwala. High-quality motion deblurring from a single image. In *ACM Transactions on Graphics (TOG)*, volume 27, page 73. ACM, 2008. 1, 7
- [28] S. Su and W. Heidrich. Rolling shutter motion deblurring. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1529–1537. IEEE, 2015. 1, 4, 7
- [29] L. Sun, S. Cho, J. Wang, and J. Hays. Edge-based blur kernel estimation using patch priors. In *Computational Photography (ICCP), 2013 IEEE International Conference on*, pages 1–8. IEEE, 2013. 1
- [30] Y.-W. Tai, X. Chen, S. Kim, S. J. Kim, F. Li, J. Yang, J. Yu, Y. Matsushita, and M. S. Brown. Nonlinear camera response functions and image deblurring: Theoretical analysis and practice. *IEEE transactions on pattern analysis and machine intelligence*, 35(10):2498–2512, 2013. 1, 2, 4, 7
- [31] Y.-W. Tai, P. Tan, and M. S. Brown. Richardson-lucy deblurring for scenes under a projective motion path. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(8):1603–1618, 2011. 1
- [32] O. Whyte, J. Sivic, and A. Zisserman. Deblurring shaken and partially saturated images. In *Proceedings of the IEEE Workshop on Color and Photometry in Computer Vision, with ICCV*, 2011. 6, 7

- [33] O. Whyte, J. Sivic, and A. Zisserman. Deblurring shaken and partially saturated images. *International Journal of Computer Vision*, 2014. 7
- [34] O. Whyte, J. Sivic, A. Zisserman, and J. Ponce. Non-uniform deblurring for shaken images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2010. 1, 2, 4, 7, 8
- [35] L. Xu and J. Jia. Two-phase kernel estimation for robust motion deblurring. In *European conference on computer vision*, pages 157–170. Springer, 2010. 1, 7
- [36] L. Xu and J. Jia. Depth-aware motion deblurring. In *Computational Photography (ICCP), 2012 IEEE International Conference on*, pages 1–8. IEEE, 2012. 1, 7
- [37] L. Xu, S. Zheng, and J. Jia. Unnatural l0 sparse representation for natural image deblurring. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1107–1114, 2013. 1, 2, 6, 7, 8
- [38] Y. Zheng, S. Nobuhara, and Y. Sheikh. Structure from motion blur in low light. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 2569–2576. IEEE, 2011. 1