Going Unconstrained with Rolling Shutter Deblurring
(Supplementary Material)

Mahesh Mohan M. R., A. N. Rajagopalan
Indian Institute of Technology Madras
{ee14d023,raju}@ee.iitm.ac.in

Gunasekaran Seetharaman
U.S. Naval Research Laboratory
guna.seetharaman@nrl.navy.mil

We begin by revisiting the problem motivation. This is followed by section S2 which contains our proof for claim 2 (section 3). Section S3 gives implementation details, and section S4 is devoted to additional evaluations.

S1. Problem Motivation (illustrative)

As mentioned in section 1 in the main paper, our method advances the state-of-the-art in RS deblurring, as it can deal with wide-angle configuration, unconstrained ego-motion and unconstrained shutter, without the need for timing information. Here, we further elaborate the significance of these problems that we have addressed in our work.

Unconstrained Focal-length: The PSFs provided in Fig. 2(b), which illustrates the importance of inplane rotations for wide-angle systems, is created using a focal length of 29 mm and real hand-held trajectory #39 in [6]. The full set of PSFs is provided in Fig. S1. We give in Fig. S2 focal-length settings of some popular CMOS imaging devices. It is clearly evident from the figure that wide-angle configurations are indeed important in photography (and predominant in cell-phones and drone cameras). However, the state-of-the-art RS-BMD [11] works only for narrow-angle settings. Hence, it is important to accommodate wide-angle settings.

Unconstrained Ego-motion: Even though a polynomial function can reasonably model human camera shake, RS blur also exists in images captured by drones, street view cars, etc., wherein the ego-motion is seldom regular [11]. Fig. S3 illustrates this fact with an under-damped response of a robotic system (which we employed in Figs. 5(d-f) using [4]). Also given is the approximation using a fourth order polynomial (as used in state-of-the-art RS-BMD [11]). From the plot it is clear that the polynomial model is unable to adequately capture the motion, thus underscoring the need for handling unconstrained ego-motion.

RS timing information: Both shutter speed ($t_e$) and inter-row delay ($t_r$) are required a priori in state-of-the-art RS-BMD [11] to fragment the motion trajectory for each image-row. Getting $t_r$ from a camera requires processing of videos taken using the same camera setting (section 5.2 in [11]). Deriving both $t_e$ and $t_r$ without the meta-data and camera information further escalates the difficulty. In contrast, our method does not need any a priori timing information. Note that we estimate the value $t_r/t_e$ for the RS prior in Eq. (9) solely from image intensities as discussed in section 4.4.

S2. Proof of Claim 2

Claim 2: The prior which restricts drifting of TSFs between blocks (in Eq. (9)) is a convex function in w, and can be represented as a norm of matrix vector multiplication, i.e., as $\|Gw\|_2^2$, with sparse G.
To prove this, we draw from the following well-known properties of convex function [1] which are a linear function is always convex (prop. 1), composition of convex functions is always convex (prop. 2), and non-negative sum of convex functions is convex (prop. 3).

Proof: Considering \( n_b \) number of image blocks and each block-MDF \( w_j \) having length \( l \), an individual additive component in our RS prior (in Eq. (9)) can be represented as 
\[
\| \Gamma (r_b(j-i+1)) \cdot S_{(i,j)} w \|_2^2,
\]
where \( S_{(i,j)} \) is a matrix of dimension \( l \times n_b \cdot l \), with all zeros except two scaled identity matrices of dimension \( l \times l \) corresponding to \( i \)th TSF (with scale 1) and \( j \)th TSF (with scale \(-1\)). Therefore, the term \( \{ \Gamma (r_b(j-i+1)) \cdot S_{(i,j)} w \} \) is a linear function in \( w \). Since \( \| \Gamma (r_b(j-i+1)) \cdot S_{(i,j)} w \|_2^2 \) is a composite of squared \( L_2 \) norm (which is convex) of a linear function in \( w \), each additive component is convex (props. 1 and 2). Resultantly, the sum of all additive components in Eq. (9), i.e., \( \text{prior}(w) \), is a convex function in \( w \) (prop. 3).

Also, \( \text{prior}(w) \) can be represented as \( \| G w \|_2^2 \), where matrix \( G \) is obtained by vertically concatenating matrices \( \{ \Gamma (r_b(j-i+1)) \cdot S_{(i,j)} \} \) corresponding to the individual additive component in RS prior. Since \( S_{(i,j)} \) is a sparse matrix, \( G \) will also be sparse. Hence proved.

### S3. Implementation Details

We implemented our algorithm in MATLAB. We empirically set 7 scales, each with 7 iterations, in our scale-space framework (section 4). The blurred image in the \( i \)th scale is formed by downsampling the input image by a factor of \( (1/\sqrt{2})^{i-1} \). To start the alternative minimization, the coarsest scale MDFs are initialized with Kronecker delta. For ego-motion estimation (section 4.2), we consistently used the RS-prior regularization (\( \alpha \) in Eq. (13)) in level \( i \) as \( 2^{7-i} \) (so that the RS prior can cope with the increasing image size, and thus the data fidelity magnitude \( \| Fw - \nabla B \|_2^2 \), in finer levels). We used the MDF regularization \( \beta' \) (in Eq. (13)) as 0.01. For latent image estimation (section 4.3), we used \( R = 48 \) such that each image-patch is square, and with 6 patches along the shorter dimension and 8 along the longer dimension. For the Richardson-Lucy deconvolution (employed in the last iteration of the finest level), we used a total number of 30 iterations. For the selection of block-size (section 4.4), we used an initial block-size \( r_0 \) as 145, and a downsampling factor of 2 (i.e., \( M_0 = M/2 \) and \( N_0 = N/2 \)).

Running time reported in Table 1 is obtained on the same system with an Intel Xeon processor with 32 GB memory. We found that for deblurring an \( 800 \times 800 \) RGB image (of maximum blur-length of 30 pixels), our unoptimized MATLAB implementation took about 9 minutes. Fig. S4 provides a detailed break-up of the time taken for each estimation step. In fact, observe that a large fraction of the total time is utilized for latent image estimation in the final iteration which involves a costly image-prior (see section 4.3). This underscores the importance of our efficient prior-less estimation in the initial iterations derived from RS-EFF (Eq. (14)).

### S4. Additional Evaluations

We provide in Fig. S5 iteration-by-iteration results to illustrate how the algorithm works. In Fig. S6, we give full images corresponding to the patches of synthetic experiment results provided in Figs. 5(a-i). In Figs. S7-S11, we give additional evaluations for the real RS-BMD examples provided in Figs. 7 & 8. These include SIV-BMD [2] and RS rectification with SIV-BMD [2] (as reported in [11]), and state-of-the-art CCD-BMD [9]. We also consider BMD without our RS prior to illustrate the ego-motion ambiguity in RS-BMD. For low-light case, we consider [5] that specifically addresses low-light BMD (albeit for CCD cameras). The codes for [5], [11] and [9] are downloaded from the author’s website and executed using default parameters. Additional examples under different lighting condition and for wide-angle settings are given in Fig. S12.

For sake of completeness; we provide GS deblurring comparisons with state-of-the-art CCD-BMD methods of [9, 8, 10, 13, 12] and [3] in Figs. S13 & S14. We evaluated on the examples from the dataset of [7] and [9] using their reported results. The results show that our method works equally well for CCD cameras and importantly, without warranting any prior knowledge of the shutter.
Figure S5. Iteration-by-iteration results of the alternative minimization of block-wise MDFs and latent image: (a-c) Estimated block-wise MDFs and (d) Estimated latent image. Notice the variation in block-wise MDFs, which depicts the characteristic of RS blur (as shown in Fig. 3). Also, observe the convergence of the block-wise MDFs through iteration 5 to 7 in the finest image scale (last three rows).
Figure S6. Full-sized images corresponding to the image patches given in Figs. 2(a-i): First row gives a case of wide-angle system (Figs. 2(a-c)), second row gives a case of vibratory motion (Figs. 2(d-f)), and third row gives a case of CCD-blur (Figs. 2(g-i)). (Best viewed on high-resolution display with zoom-in corresponding to an $800 \times 800$ image size.)
Figure S7. Detailed comparisons for RS narrow-angle example in dataset [11] (Fig. 7-top-row). Note the effect of incoherent combination due to the block shift-ambiguity (section 3, claim 1) in (i)-first row, which is successfully suppressed by our RS prior ((i)-second row).

Figure S8. Detailed comparisons for RS narrow-angle example in dataset [11] (Fig. 7-second-row). Our method recovers finer details (see bag-zipper in patch 1), and deblur with negligible ringing artefacts (see bag-badge in patch 2), as compared to competing methods.
Figure S9. Detailed comparisons for RS wide-angle example (Fig. 8-first row). In contrast to competing methods, our method models the RS ego-motion better (observe the residual blur in the letters, and the repeated occurrence of the longest grass leaf in (c)).

Figure S10. Comparisons for RS wide-angle case (Fig. 8-second row). White boxes in images (e) and (f) show the effect of RS prior.

Figure S11. Comparisons for RS wide-angle example (Fig. 8-third row). White box in images (e) and (f) shows the effect of RS prior.
Figure S12. Additional RS comparisons with state-of-the-art RS-BMD method [11] under different lighting conditions and for wide-angle settings. Note the inefficacy of the competing method in dealing with wide angle systems.

Figure S13. Comparisons for CCD blur example in dataset [7]. Our result is comparable with [8, 10, 13] and [9].

Figure S14. Comparisons for CCD blur example in dataset [9]. Our result is comparable with [3, 12, 13] and [9].
References


