SEMI-BLIND IMAGE DE-BLURRING WITH IMU PRIOR

Haiyan Shu and Zhengguo Li

SRO Department, Institute for Infocomm Research, 1 Fusionopolis Way, Singapore

ABSTRACT

Due to fast movement of an autonomous system, motion blur is an issue for images with long-exposures, especially for those captured at dim-lighting conditions. In this paper, a simple semi-blind image de-blurring algorithm is designed by using measurements from an inertial measurement unit (IMU) embedded in the autonomous system. A prior of the blur kernel is first derived from an elegant IMU integration model. The image is then de-blurred by using the prior of the blur kernel with a refinement step to further mitigate the drift caused by the inaccurate IMU measurements. Experimental results show the improvement of the proposed algorithm for image de-blurring.

1. INTRODUCTION

Increasing amounts of attention have been received in the autonomous driving market [1, 2, 3, 4]. Visual information is crucial for the autonomous systems which require high fidelity visual information captured and/or reconstructed. However, due to the fast movement, motion blur caused by long-exposure photography is unavoidable in the image capturing process [5, 6], especially in low-lighting conditions. Motion de-blurring becomes an important topic for the autonomous systems.

The model-based blind de-blurring is a popular topic where no motion information is available [7, 8, 9]. In these approaches, maximum a posterior (MAP) and priors are developed to derive the blur kernel. Iteration steps are then applied to refine the de-blurring operations. Later, deep learning based approaches are employed to solve this problem [10, 11]. Non-blind de-blurring is also developed by utilizing motion information from additional sensors. The motion information is adopted to derive the camera motion and the corresponding blur kernel and thus restore the original image. Secondary detector was used in [12] to capture a sequence of images. The motion of the camera can then be derived from these fixed intervals during the exposure time. Inertial measurement unit (IMU) was adopted in [13] to provide an accurate measurement of the motion. Drift compensation is addressed by searching the potential moving direction over the space with high computational demanded. De-convolution scheme with embedded deep learning component is also introduced which can be employed in non-blind motion de-blurring [14]. However, the measurement of IMU reading may not be accurate enough to address the blurring process. Further process is necessary to enhance the visual output.

In this paper, a novel semi-blind image de-blurring algorithm is proposed for the autonomous systems with IMUs by using model-based deep learning. A prior on the blur kernel is first derived by using an elegant IMU integration model [15]. Same as the algorithm in [13], all pixels in the image are assumed to have the same depth. Our mathematical derivations show that the transformation matrix between two images is spatially invariant. Subsequently, the blur kernel is spatially invariant. This is different from the observation on the blur kernel in [13]. A non-blind algorithm is obtained by combining the algorithm in [14] and the prior on the blur kernel, and it is applied to initially reduce the motion blur. There is drift on the blur kernel caused by the IMU noise [16]. Inspired by the existing blind de-blurring algorithms [7, 8, 9], a quadratic optimization problem is formulated to refine the blur kernel. It is much simpler to solve the quadratic optimization problem than the corresponding problem on drift reduction in [13]. The blurred image is finally restored by the refined blur kernel. The overall de-blurring algorithm is a semi-blind one. It is worth noting that the resultant non-blind and semi-blind de-blurring algorithms are model-based deep learning methods [17]. Experimental results show that the prior from the IMU measurement is indeed very useful for the reduction of motion blue, and the semi-blind algorithm outperforms the non-blind one due to the drift reduction.

The rest of this paper is organized as follows. The prior of blur kernel is derived in Sec. 2. A semi-blind de-blurring algorithm is proposed in Sec. 3. Experimental results are presented in Sec. 4. Finally, conclusion remarks are given in Sec. 5.

2. PRIOR OF THE BLUR KERNEL

In this section, a prior of the blur kernel is derived by using the IMU motion integration model in [15] under an assumption that all the pixels in an image have the same depth.

2.1. IMU Integration of 3D Points

An IMU motion integration model is provided for 3D points of stationary objects in the body frame [15] with the notations in Table 1. The body frame is the camera frame rather than the IMU frame. a_b and ω_b are computed by using the corresponding IMU measurements and rotation matrix as $R_{bi}a_i$ and $R_{bi}\omega_i$, respectively. Clearly, both the a_b and ω_b are also constant between two successive IMU measurements.

The dynamics of the lth 3D point (belonging to a stationary object) in the body frame is represented as [15]

$$\begin{bmatrix} \dot{P}_{b,l} \\ \dot{V}_b \\ \dot{R}_1 \\ \dot{R}_2 \\ \dot{R}_3 \end{bmatrix} = A(\omega_b) \begin{bmatrix} P_{b,l} \\ V_b \\ R_1 \\ R_2 \\ R_3 \end{bmatrix} + \begin{bmatrix} 0 \\ a_b \\ 0 \\ 0 \\ 0 \end{bmatrix}, \tag{1}$$

Table 1. Notations used in the IMU integration model.

$V_b = \begin{bmatrix} V_x & V_y & V_z \end{bmatrix}^T$	linear velocity in body frame
$R_{bw} = \begin{bmatrix} R_1 & R_2 & R_3 \end{bmatrix}$	rotation matrix from world frame to body frame
$a_b = \begin{bmatrix} a_x & a_y & a_z \end{bmatrix}^T$	linear acceleration in body frame
a_i	linear acceleration in IMU frame
$g_w = \begin{bmatrix} 0 & 0 & g \end{bmatrix}^T$	gravitational vector
$\omega_b = \begin{bmatrix} \omega_x & \omega_y & \omega_z \end{bmatrix}^T$	angular velocity in body frame
ω_i	angular velocity in IMU frame
R_{ib}	rotation matrix from IMU frame to body frame
$P_{b,l} = \begin{bmatrix} X_l & Y_l & Z_l \end{bmatrix}^T$	coordinate of the lth 3D point in body frame
m_l=$\begin{bmatrix} u_l & v_l \end{bmatrix}^T$	coordinate of the lth 2D pixel in an image
ς	duration of interval between two successive IMU measurements

where the matrix $A(\omega_b)$ is given by

$$A(\omega_b) = -\begin{bmatrix} \omega_b^\wedge & I & 0 & 0 & 0 \\ 0 & \omega_b^\wedge & 0 & 0 & -gI \\ 0 & 0 & \omega_b^\wedge & 0 & 0 \\ 0 & 0 & 0 & \omega_b^\wedge & 0 \\ 0 & 0 & 0 & 0 & \omega_b^\wedge \end{bmatrix}, \quad (2)$$

I is a 3×3 identity matrix, and $\omega_b^\wedge$ is defined as

$$\omega_b^\wedge = \begin{bmatrix} 0 & -\omega_z & \omega_y \\ \omega_z & 0 & -\omega_x \\ -\omega_y & \omega_x & 0 \end{bmatrix}. \quad (3)$$

Since both ω_b and a_b are constant in an interval between any two successive IMU measurements, the system (1) is a time invariant linear system in the interval. A discrete motion model can be derived for the lth 3D point as [15]

$$\begin{cases} P_{b,l}(k+1) = E(-\theta_b(k))P_{b,l}(k) - \varsigma E(-\theta_b(k))V_b(k) - \\ \qquad \frac{g}{2}\varsigma^2 E(-\theta_b(k))R_3(k) - \Upsilon(-\theta_b(k))a_b(k)\varsigma^2 \\ V_b(k+1) = E(-\theta_b(k))V_b(k) + g\varsigma E(-\theta_b(k))R_3(k) \\ \qquad + J_r(-\theta_b(k))a_b(k)\varsigma \\ R_{bw}(k+1) = E(-\theta_b(k))R_{bw}(k) \end{cases},$$

where θ_b is $\omega_b\varsigma$. The matrices $E(\theta_b)$, $J_r(\theta_b)$ and $\Upsilon(\theta_b)$ are

$$\begin{cases} E(\theta_b) = I + \frac{s}{\|\theta_b\|}\theta_b^\wedge + \frac{1-c}{\|\theta_b\|^2}(\theta_b^\wedge)^2 \\ J_r(\theta_b) = I + \frac{1-c}{\|\theta_b\|^2}\theta_b^\wedge + \frac{\|\theta_b\|-s}{\|\theta_b\|^3}(\theta_b^\wedge)^2 \\ \Upsilon(\theta_b) = \frac{I}{2} + \frac{s-\|\theta_b\|c}{\|\theta_b\|^3}\theta_b^\wedge + \frac{1-c+\frac{\|\theta_b\|^2}{2}}{\|\theta_b\|^4} - \frac{\|\theta_b\|s}{\|\theta_b\|^4}(\theta_b^\wedge)^2 \end{cases}, \quad (4)$$

and $c = \cos\|\theta_b\|$ and $s = \sin\|\theta_b\|$.

$\omega_b(k)$ and $a_b(k)$ are usually different from $\omega_b(k+1)$ and $a_b(k+1)$. Therefore, the dynamics of the 3D point is a switched system between the two time instances [18, 19]. By accumulating k from i to $(j-1)$, the discrete motion pre-integration model between two time instances of $k = i$ and $k = j$ can be derived for the lth 3D point as

$$\begin{cases} P_{b,l}(j) = F^T(i,j)P_{b,l}(i) + t(i,j) \\ V_b(j) = F^T(i,j)(V_b(i) + g\sum_{k=i}^{j-1}\varsigma R_3(i) + \mu_b(i,j)) \\ R_{bw}(j) = F^T(i,j)R_{bw}(i) \end{cases}, \quad (5)$$

where $t(i,j)$ is a translation vector, $F(i,j)$ is a rotation matrix, they and $\mu_b(i,j)$ are given as

$$\begin{cases} t(i,j) = -F^T(i,j)(\chi_b(i,j) + \zeta_b(i,j)) \\ F(i,j) = \prod_{k=i}^{j-1} E(\theta_b(k)) \\ \mu_b(i,j) = \sum_{k=i}^{j-1} F(i,k)J_r(\theta_b(k))a_b(k)\varsigma \end{cases}, \quad (6)$$

$F(k,k) = I$ and $\mu_b(k,k) = 0$ for all k's, and the $\chi_b(i,j)$, $\zeta_b(i,j)$, and $\Lambda(\theta_b)$ are

$$\begin{cases} \chi_b(i,j) = V_b(i)\sum_{k=i}^{j-1}\varsigma + \frac{g}{2}(\sum_{k=i}^{j-1}\varsigma)^2 R_3(i) \\ \zeta_b(i,j) = \sum_{k=i}^{j-1}(F(i,k)\Lambda(\theta_b(k))a_b(k)\varsigma^2 + \mu_b(i,k)\varsigma) \\ \Lambda(\theta_b) = \frac{1}{2}I + \frac{\theta_b - s}{\|\theta_b\|^3}\theta_b^\wedge + \frac{2c-2+\|\theta_b\|^2}{2\|\theta_b\|^4}(\theta_b^\wedge)^2 \end{cases}, \quad (7)$$

Let $b^a(i)$ and $b^\omega(i)$ be the biases of a_b and ω_b in the interval $[i\varsigma, j\varsigma]$. The values of $a_b(k)$ and $\omega_b(k)$ will be replaced by $(a_b(k) - b^a(i))$ and $(\omega_b(k) - b^\omega(i))$, respectively [16]. The transformation matrix from $P_{b,l}(i)$ to $P_{b,l}(j)$ is

$$T(i,j) = \begin{bmatrix} F^T(i,j) & t(i,j) \\ 0 & 1 \end{bmatrix}, \quad (8)$$

and it is spatially invariant.

2.2. Motion Prior from the IMU Integration

In this subsection, a prior of the blur kernel is derived by using the motion integration model (5). The relationship between the lth 2D pixel m_l in an image and the 3D point $P_{b,l}$ for a look-front camera is represented as

$$X_{b,l}\begin{bmatrix} u_l \\ v_l \\ 1 \end{bmatrix} = \begin{bmatrix} f_u & 0 & c_u \\ 0 & f_v & c_v \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} Y_{b,l} \\ Z_{b,l} \\ X_{b,l} \end{bmatrix}, \quad (9)$$

where f_u and f_v are the focal lengths of the camera in pixels, and c_u and c_v are the coordinates of the principle point in pixels.

It can be derived from the equation (8) that

$$\begin{bmatrix} X_{b,l}(j) \\ Y_{b,l}(j) \\ Z_{b,l}(j) \end{bmatrix} = \begin{bmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{bmatrix}\begin{bmatrix} X_{b,l}(i) \\ Y_{b,l}(i) \\ Z_{b,l}(i) \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}. \quad (10)$$

From the pin-hole camera model (9), it can be known that

$$\begin{bmatrix} u_l(k) \\ v_l(k) \end{bmatrix} = \begin{bmatrix} f_u\frac{Y_{b,l}(k)}{X_{b,l}(k)} + c_u \\ f_v\frac{Z_{b,l}(k)}{X_{b,l}(k)} + c_v \end{bmatrix}; \; k \in \{i,j\}. \quad (11)$$

It can be derived from the equation (10) that

$$\begin{cases} Y_{b,l}(j) = F_{21}X_{b,l}(i) + F_{22}Y_{b,l}(i) + F_{23}Z_{b,l}(i) + t_y \\ Z_{b,l}(j) = F_{31}X_{b,l}(i) + F_{32}Y_{b,l}(i) + F_{33}Z_{b,l}(i) + t_z \end{cases}.$$

Defining a matrix $\hat{F}$ as

$$\hat{F} = \frac{X_{b,l}(i)}{X_{b,l}(j)} \begin{bmatrix} F_{22} & \frac{f_u F_{23}}{f_v} \\ \frac{f_v F_{32}}{f_u} & F_{33} \end{bmatrix} \tag{12}$$

and a vector $\hat{t}$ as

$$\hat{t} = \begin{bmatrix} c_u \\ c_v \end{bmatrix} + \frac{X_{b,l}(i)}{X_{b,l}(j)} \left(\begin{bmatrix} f_u F_{21} \\ f_v F_{31} \end{bmatrix} - \hat{F} \begin{bmatrix} c_u \\ c_v \end{bmatrix} \right) + \frac{1}{X_{b,l}(j)} \begin{bmatrix} f_u t_y \\ f_v t_z \end{bmatrix}, \tag{13}$$

it follows that

$$\begin{bmatrix} u_l(j) \\ v_l(j) \end{bmatrix} = \hat{F} \begin{bmatrix} u_l(i) \\ v_l(i) \end{bmatrix} + \hat{t}. \tag{14}$$

Clearly, both the matrix $\hat{F}$ and the vector $\hat{t}$ are independent of the 3D points *if and only if* all the 3D points have the same depth. All the 3D points are assumed to have the same depth in [13]. Therefore, the transformation matrix between two images is not spatially varying as indicated in [13] but spatially invariant. The transformation (14) becomes an affine transformation which is widely used in visual deepfake [20]. For simplicity, the proposed algorithm is on top of the same assumption in [13]. Instead of using the complicated algorithm in [13], the algorithm in [12] is extended to consider the motion model (14).

Same as the algorithm in [12], the blur kernel is defined as κ: $(u, v) \mapsto e$ with e be the energy level at the position (u, v). The function κ satisfies

$$\int\int \kappa(u, v) du dv = 1, \tag{15}$$

where the function $\kappa(u, v)$ is implemented via a pair of path function f: $t \mapsto (u, v)$ and an energy function h: $t \mapsto e$. Here, $f(t)$ is continuous and at least twice differentiable. It is obtained by using the spline interpolation on the discrete motion samples from the equation (14). And the function $h(t)$ is required to satisfy

$$\int_t^{t+\delta t} h(\tau) d\tau = \frac{\delta t}{t_e - t_s},\ t_s \le t \le t_e - \delta t, \tag{16}$$

where $[t_s, t_e]$ is the long-exposure interval.

Similar to the approach in [12], the energy function h: $t \mapsto e$ is derived as the inverse velocity at the position (u, v) at corresponding time t. This is based on the assumption that the higher velocity measured, the lower energy captured at the corresponding time interval. This velocity information can be easily derived from IMU measurement with higher accuracy than the algorithm in [12] where the position interpolation was used. In the velocity based approach, the path $f(t)$ is spitted into intervals defined by the IMU measurement interval (ς). The corresponding energy function $h(t)$ is estimated by 2D velocity $V_u(t)$ and $V_v(t)$:

$$h(t) = \frac{1}{\sqrt{V_u^2(t) + V_v^2(t)} \cdot \varsigma}. \tag{17}$$

where 2D velocity can be derived from 3D velocity given as

$$\begin{bmatrix} V_u(k) \\ V_v(k) \end{bmatrix} = \begin{bmatrix} f_u \left(\frac{V_y(k)}{X_{b,l}(k)} - \frac{Y_{b,l}(k) V_x(k)}{X_{b,l}^2(k)} \right) \\ f_v \left(\frac{V_z(k)}{X_{b,l}(k)} - \frac{Z_{b,l}(k) V_x(k)}{X_{b,l}^2(k)} \right) \end{bmatrix}. \tag{18}$$

This velocity based scheme usually presents higher accuracy than the position based approach. Smoothing and normalization are then applied to make sure the energy conservation constraint (16) is satisfied. The prior on the blur kernel κ can be finally derived from the estimated functions $f(t)$ and $h(t)$. For simplicity, the prior is denoted as κ_0.

3. SEMI-BLIND IMAGE DE-BLURRING

The common image blurring model is given by

$$y = \kappa \otimes x + \eta, \tag{19}$$

where y is the the blurred image which is assumed to be the convolution of a clear image x with a blur kernel κ, and η is the Gaussian noise.

When the blur kernel κ is available, the classical solution to restore the clear image is derived by the Wiener filter [21]

$$\hat{x} = \mathcal{F}^{-1} \left(\frac{\mathcal{F}(\kappa \odot y)}{|\mathcal{F}(\kappa)|^2 + n/S} \right), \tag{20}$$

where $\mathcal{F}$ is the two-dimensional discrete Fourier transform, S and n are the corresponding image and noise's expected power spectra, and $\odot$ is the entry-wise product.

With the derived prior on the blur kernel κ_0 in the previous section, the non-blind image de-blurring can be adopted to restore the blurred image. Most of existing non-blind de-blurring algorithms from Richardson-Lucy deconvolution [22, 23] to fast Fourier transform (FFT) based approach failed to address image boundary handling, even with an "edgetaper" operation [24] to apply additional circular convolution to the boundary region. In [14], a convolutional neural network (CNN) is incorporated into the FFT based solution with a novel boundary adjustment method to alleviate the problematic circular convolution assumption. This approach introduces a specialized CNN based term $\phi_t^{CNN}(x^t)$ for every step of iteration and simultaneously utilizes the efficiency of FFT-based approach. This derivation is given as

$$x^{t+1} = \mathcal{F}^{-1} \left(\frac{\mathcal{F}(\kappa_0 \odot y + \frac{1}{\omega_t(\lambda)} \phi_t^{CNN}(x^t))}{|\mathcal{F}(\kappa_0)|^2 + \frac{1}{\omega_t(\lambda)} \sum_i |\mathcal{F}(f_{it})|^2} \right), \tag{21}$$

where $\omega_t(\lambda)$ is a learned scalar function. A non-blind de-blurring algorithm is available by combining the equation (21) and the derived blur kernel in the section 2.2.

The blur kernel is assumed to be accurate by the non-blind image de-blurring algorithm (21). However, there exists a deviation between the actual IMU measurements and the groundtruth [13]. This deviation of the computed motion from the true motion makes the kernel estimation inaccurate for the subsequent de-blurring operation. A compensation approach is adopted in [13] to search in a small local neighborhood around the initially computed end point via an energy minimization framework. Different from [13], the proposed semi-blind de-blurring algorithm adopts a simple refine-

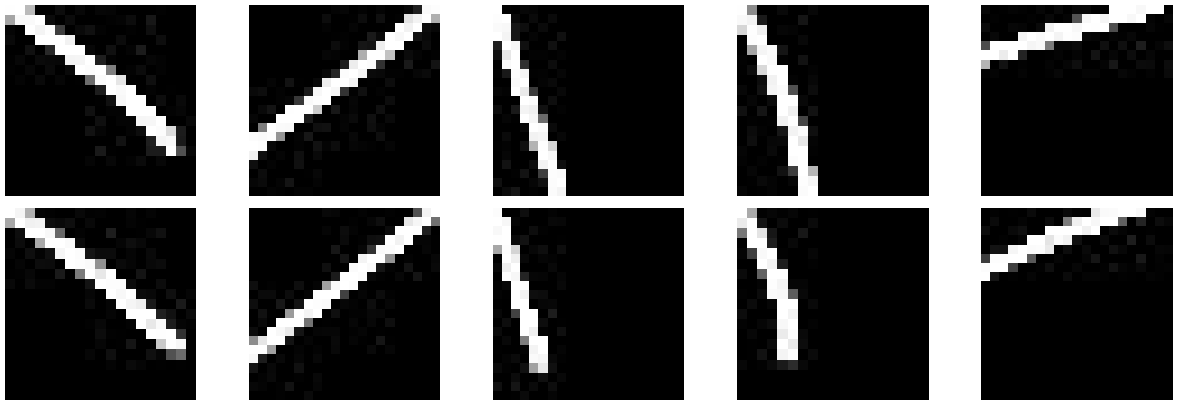


Fig. 1. Illustration of blur kernels. Top: groundtruth; Bottom: Priors from IMU measurements.

ment step on the blur kernel κ as

$$\min_{\kappa}\{\|\nabla\hat{x}\otimes\kappa-\nabla y\|_2^2+\gamma\|\kappa-\kappa_0\|_2^2\}, \tag{22}$$

where ∇ is the gradient operation.

Our refinement is inspired by the blind de-blurring algorithm [7]. The kernel refinement is a least square problem which can be solved efficiently using the FFT. After computing the κ, necessary non-negative and normalization operations are applied to get the refined κ so as to satisfy the requirement (15). With the refinement step, possible deviation of the IMU measurement from the ground-truth can be mitigated. The de-blurring operation can be conducted again with the refined κ to improve the performance. Due to the CNN based term $\phi_t^{CNN}(x^t)$ in the equation (21), the non-blind and semi-blind de-blurring algorithms in this section are model-based deep learning methods [17].

4. EXPERIMENTAL RESULTS

Prior of the blur kernel in the section 2.2 is verified by comparing the non-blind and semi-blind de-blurring algorithms in the section 3 with the blind de-blurring approach on top of the dark channel prior (DCP) [8]. Two popular datasets are tested, one is from [25] which includes four low quality images and the other from [26] where images are from five groups including manmade, natural, people, saturated, and text.

We employ the EuRoC micro aerial vehicle (MAV) Dataset to compute the blur kernel. The IMU measurements are collected using an MAV in different places with different conditions in flight dynamics [27]. We choose the IMU information from "Dataset Vicon room 1 difficult". In this dataset, the IMU measurements were sampled at 200Hz and a Leica Nova MS503 laser tracker was used to provide the measurement for groundtruth reference. As such, there are two measurements at the same time, one from groundtruth and one from IMU reading. The corresponding blur kernels are then derived from these two measurements. In the simulation, 5 sets of blur kernels are generated and illustrated in Fig. 1, where the tops are from the groundtruth and the bottoms are from IMU measurements. For kernel sets **A** and **B**, the IMU measurements are close to those of the groundtruth. For kernel sets **C**, **D** and **E**, the IMU measurements deviate from the groundtruth.

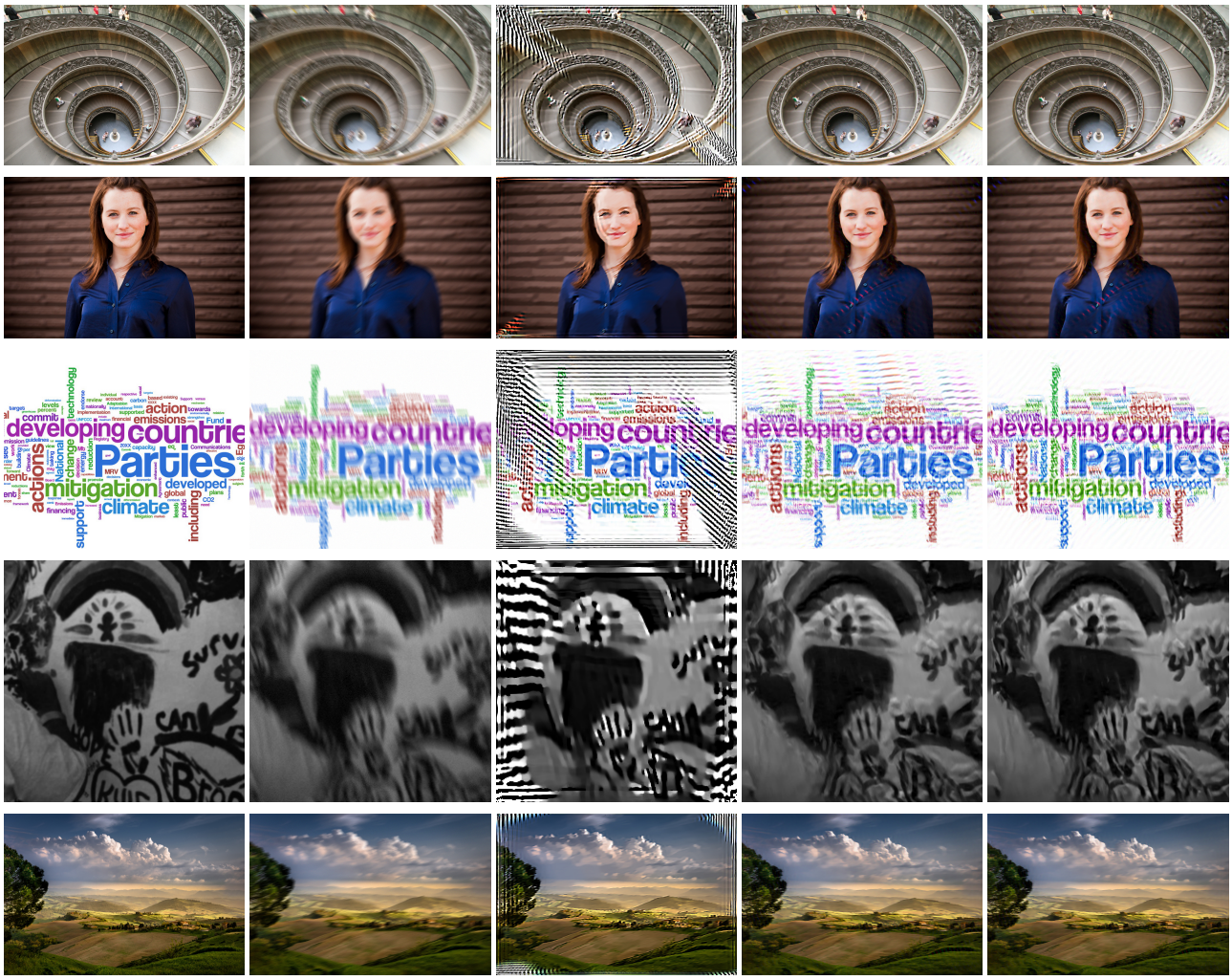
The blurred images are generated by the groundtruth kernels with the standard deviation of Gaussian noise at 1% of the dynamic range. The comparison results of PSNR are presented in Table 2. Obviously, the non-blind and semi-blind de-bluring approaches in the section 3 significantly outperform the DCP which is a blind de-blurring method. This implies that the prior from the IMU motion integration is indeed very useful for the image de-blurring.

When the estimated kernel is close to the ground-truth, there is a limited space for improvement by the semi-blind de-blurring. For the kernel sets **A** and **B**, the resultant semi-blind approach can achieve a PSNR improvement at $0.01dB$ in average. Some examples are illustrated in Fig. 2. It is also very difficult to perceive the difference between the non-blind and semi-blind deblurred images in the top two rows of Fig. 2. Therefore, the refinement can be omitted if the prior of motion blur kernel is accurate.

When the kernel deviates from the ground-truth, there are more space for the improvements. For the kernel sets **C**, **D** and **E**, average PSNR improvements of $0.185dB$, $0.098dB$, and $0.079dB$ are observed respectively. It can be shown from the bottom three rows that the delburred images by the semi-blind algorithm are sharper than the deblurred images by the non-blind algorithm on top of inaccurate blur kernels. Thus, the refinement is necessary if the prior of motion blur kernel is inaccurate.

The comparison results of structural similarity index measure (SSIM) are also presented in the same table. It also shows that the improvement is consistent by the semi-blind algorithm with respect to the non-blind algorithm due to the refinement step. These experimental results imply that it is important to derive the prior on the blur kernel as accurately as possible. As such, the refinement can be omitted and the complexity can be reduced significantly.

It is also found from Fig. 2 that, even with the "edgetaper" operation, the DCP approach cannot handle the image boundary well while the non-blind and semi-blind algorithms in the section 3 handle image boundary well due to the algorithm in [14]. On the other hand, since the algorithm in [14] is on top of the CNN, fine details could be lost in the de-blurred image when the blur kernel type from the real scenario is not covered by the training data, especially when the blur kernel is inaccurate. Some improvement in details can be observed from the semi-blind algorithm over the non-blind de-blurring approach. Further improvement can be achieved via a conventional detail enhancement algorithm [28].



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

Fig. 2. Comparison of different de-blurring algorithms. (a) groundtruth images, and (b) blurred images, as well as de-blurred images by (c) DCP, (d) proposed non-blind one, and (e) proposed semi-blind one.

Table 2. Comparison of PSNR and SSIM

kernel	PSNR (dB)			SSIM		
	DCP	non-blind	semi-blind	DCP	non-blind	semi-blind
A	29.512	31.489	31.502	0.335	0.737	0.740
B	29.396	31.007	31.018	0.303	0.553	0.557
C	29.268	30.208	30.392	0.250	0.546	0.587
D	29.528	30.570	30.667	0.296	0.518	0.543
E	29.372	30.783	30.861	0.282	0.626	0.634

5. CONCLUSION REMARKS

A simple semi-blind image de-blurring has been provided in this paper by using an elegant IMU motion integration model to derive a prior on the blur kernel. The blurred image is restored by using the prior of the blur kernel with the consideration of drift in the IMU measurement. The proposed motion de-blurring approach presents a better result with the fusion of IMU information and a refinement step.

One important application of the proposed algorithm is to reduce motion blue for differently exposed stereo-imaging. Unlike the conventional stereo imaging [29], each pair of images has two different exposures [5, 6] which can synthesize 3D high dynamic range images without ghosting artifacts [30] for augmented reality and Metaverse. This proposed algorithm can be applied to reduce the motion blur of the long-exposed image. This problem will be studied in our future research.

References

[1] M. Simsek, A. Aijaz, M. Dohler, J. Sachs, and G. Fettweis, "5G-enabled tactile internet," *IEEE Journal on Selected*

Areas in Communications, vol. 34, no. 3, pp. 460–473, 2016.

[2] V. P. Srini, "A vision for supporting autonomous navigation in urban environments," *Computer*, vol. 39, no. 12, pp. 68–77, 2006.

[3] L. Ruotsalainen, V. Renaudin, L. Pei, M. Piras, J. Marais, E. Cavalheri, and S. Kaasalainen, "Toward autonomous driving in arctic areas," *IEEE Intelligent Transportation Systems Magazine*, vol. 12, no. 3, pp. 10–24, 2020.

[4] D. Tagliaferri, M. Rizzi, M. Nicoli, S. Tebaldini, I. Russo, A. V. Monti-Guarnieri, C. M. Prati, and U. Spagnolini, "Navigation-aided automotive sar for high-resolution imaging of driving environments," *IEEE Access*, vol. 9, pp. 35 599–35 615, 2021.

[5] F. Kou, Z. Wei, W. Chen, X. Wu, C. Wen, and Z. Li, "Intelligent detail enhancement for exposure fusion," IEEE Trans. on Multimedia, vol. 20, no. 2, pp. 484–495, Feb. 2018.

[6] Y. Yang, W. Cao, S. Wu, and Z. Li, "Multi-scale fusion of two large-exposure-ratio images," *IEEE Signal Processing Letters*, vol. 25, no. 12, pp. 1885–1889, 2018.

[7] S. Cho and S. Lee, "Fast motion deblurring," in *ACM SIGGRAPH Asia 2009 papers*, 2009, pp. 1–8.

[8] J. Pan, D. Sun, H. Pfister, and M.-H. Yang, "Blind image deblurring using dark channel prior," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 1628–1636.

[9] Y. Yan, W. Ren, Y. Guo, R. Wang, and X. Cao, "Image deblurring via extreme channels prior," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 4003–4011.

[10] O. Kupyn, V. Budzan, M. Mykhailych, D. Mishkin, and J. Matas, "Deblurgan: Blind motion deblurring using conditional adversarial networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 8183–8192.

[11] D. Ren, K. Zhang, Q. Wang, Q. Hu, and W. Zuo, "Neural blind deconvolution using deep priors," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 3341–3350.

[12] S. K. Nayar and M. Ben-Ezra, "Motion-based motion deblurring," *IEEE transactions on pattern analysis and machine intelligence*, vol. 26, no. 6, pp. 689–698, 2004.

[13] N. Joshi, S. B. Kang, C. L. Zitnick, and R. Szeliski, "Image deblurring using inertial measurement sensors," *ACM Transactions on Graphics (TOG)*, vol. 29, no. 4, pp. 1–9, 2010.

[14] J. Kruse, C. Rother, and U. Schmidt, "Learning to push the limits of efficient fft-based image deconvolution," in *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 4586–4594.

[15] J. Henawy, Z. Li, W. Y. Yau, G. Seet, and K. W. Wan, "Piecewise linear de-skewing for lidar inertial odometry," in *IEEE Intelligent Transportation Systems Conference*, 2021.

[16] J. Henawy, Z. Li, W. Y. Yau, and G. Seet, "Accurate imu factor using switched linear systems for vio," *IEEE Transactions on Industrial Electronics*, vol. 68, no. 8, pp. 7199–7208, 2020.

[17] N. Shlezinger, J. Whang, Y. C. Eldar, and A. G. Dimakis, "Model-based deep learning," *arXiv preprint arXiv:2012.08405*, 2020.

[18] Z. Li, Y. C. Soh, and C. Wen, "Robust stability of a class of hybrid nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 46, no. 6, pp. 897–903, 2001.

[19] Z. Li, Y. Soh, and C. Wen, "Switched and impulsive systems: Analysis, design and applications," *Springer Science & Business Media*, 2005, vol. 313.

[20] B. Dolhansky, J. Bitton, B. Pflaum, J. Lu, R. Howes, M. Wang, and C. Canton Ferrer, "The deepfake detection challenge dataset," *arXiv e-prints*, pp. arXiv–2006, 2020.

[21] N. Wiener et al., "Extrapolation, interpolation, and smoothing of stationary time series: with engineering applications," *MIT press Cambridge, MA*, 1964, vol. 8.

[22] W. H. Richardson, "Bayesian-based iterative method of image restoration," *JoSA*, vol. 62, no. 1, pp. 55–59, 1972.

[23] L. B. Lucy, "An iterative technique for the rectification of observed distributions," *The astronomical journal*, vol. 79, p. 745, 1974.

[24] S. J. Reeves, "Fast image restoration without boundary artifacts," *IEEE Transactions on image processing*, vol. 14, no. 10, pp. 1448–1453, 2005.

[25] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman, "Understanding and evaluating blind deconvolution algorithms," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 1964–1971.

[26] W. S. Lai, J. B. Huang, Z. Hu, N. Ahuja, and M.-H. Yang, "A comparative study for single image blind deblurring," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 1701–1709.

[27] M. Burri, J. Nikolic, P. Gohl, T. Schneider, J. Rehder, S. Omari, M. W. Achtelik, and R. Siegwart, "The Euroc micro aerial vehicle datasets," *The International Journal of Robotics Research*, vol. 35, no. 10, pp. 1157–1163, 2016.

[28] F. Kou, W. Chen, Z. Li, and C. Wen, "Content adaptive image detail enhancement," IEEE Signal Processing Letters, vol. 22, no. 2, pp. 211-215, Feb. 2015.

[29] H. Hirschmuller, "Stereo processing by semiglobal matching and mutual information," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 30, no. 2, pp. 328–341, 2007.

[30] J. Zheng, Z. Li, Z. Zhu, S. Wu, and S. Rahardja, "Hybrid patching for a sequence of differently exposed images with moving objects," *IEEE transactions on image processing*, vol. 22, no. 12, pp. 5190–5201, 2013.