Image as an IMU: Estimating Camera Motion from a Single Motion-Blurred Image

Supplementary Material

1. Pixel Motion Derivation

Suppose $\mathbf{P} \in \mathbb{R}^3$ is a 3D point in space, f is the focal length of the camera, and d is the depth. Given the standard pinhole camera perspective projection equation:

$$\mathbf{p} = f\frac{\mathbf{P}}{d}, \tag{1}$$

Trucco and Verri [1] derive Equations 2, 3 relating pixel velocity and the camera motion by taking the time derivative of both sides of Equation 1:

$$\begin{aligned}\dot{p}_x = &\frac{v_z p_{x,1} - v_x f}{d} - \omega_y f + \omega_z p_{y,1} \\ &+ \frac{\omega_x p_{x,1} p_{y,1}}{f} - \frac{\omega_y (p_{x,1})^2}{f}\end{aligned} \tag{2}$$

$$\begin{aligned}\dot{p}_y = &\frac{v_z p_{y,1} - v_y f}{d} + \omega_x f - \omega_z p_{x,1} \\ &- \frac{\omega_y p_{x,1} p_{y,1}}{f} + \frac{\omega_x (p_{y,1})^2}{f}.\end{aligned} \tag{3}$$

We assume that the velocities $\mathbf{v}$ and $\boldsymbol{\omega}$ are constant across exposure. Therefore, we multiply each side by τ, the exposure time, and reformulate the pixel velocity equations:

$$\begin{aligned}F_x = &\frac{t_z p_x - t_x f}{d} - \theta_y f + \theta_z p_y \\ &+ \frac{\theta_x p_x p_y}{f} - \frac{\theta_y (p_x)^2}{f}\end{aligned}$$

$$\begin{aligned}F_y = &\frac{t_z p_y - t_y f}{d} + \theta_x f - \theta_z p_x \\ &- \frac{\theta_y p_x p_y}{f} + \frac{\theta_x (p_y)^2}{f}.\end{aligned}$$

2. Evaluation Sequences.

We include several frames of each evaluation sequence to demonstrate the severity of the motion blur. These can be seen in Figure 1.

3. Additional Velocity Predictions.

Figures 2-4 show the velocity prediction comparisons for the rest of the video sequences: commonroom, dining, and office, respectively. Since COLMAP (DISK+LightGlue) failed for the dining sequence, we show the output from DROID-SLAM.

References

[1] Emanuele Trucco and Alessandro Verri. *Introductory Techniques for 3-D Computer Vision*, pages 178–184. Prentice Hall PTR, USA, 1998. 1

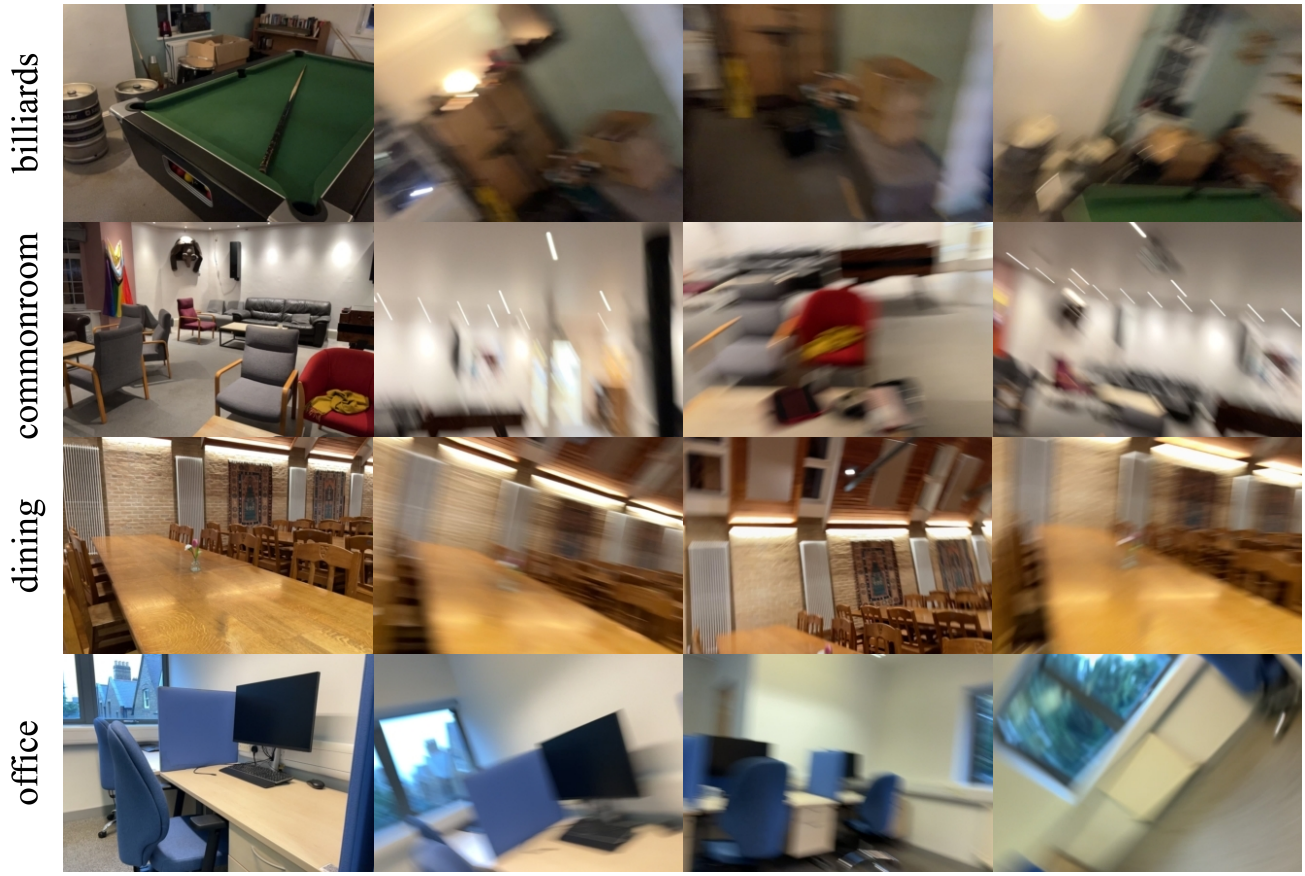

Figure 1. Frames that highlight the motion blur extent in the evaluation sequences.

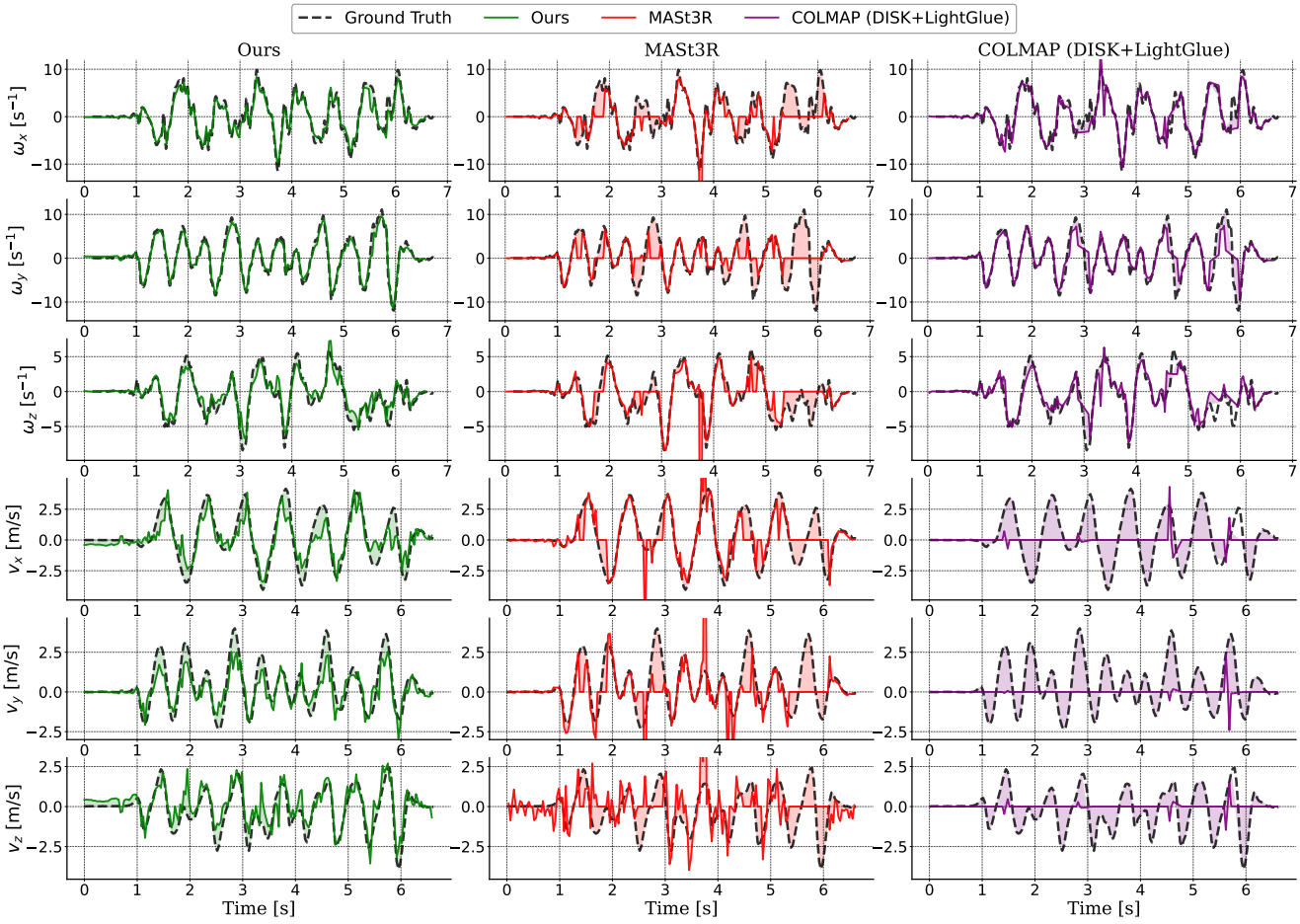

Figure 2. Velocity prediction comparison for the commonroom sequence.

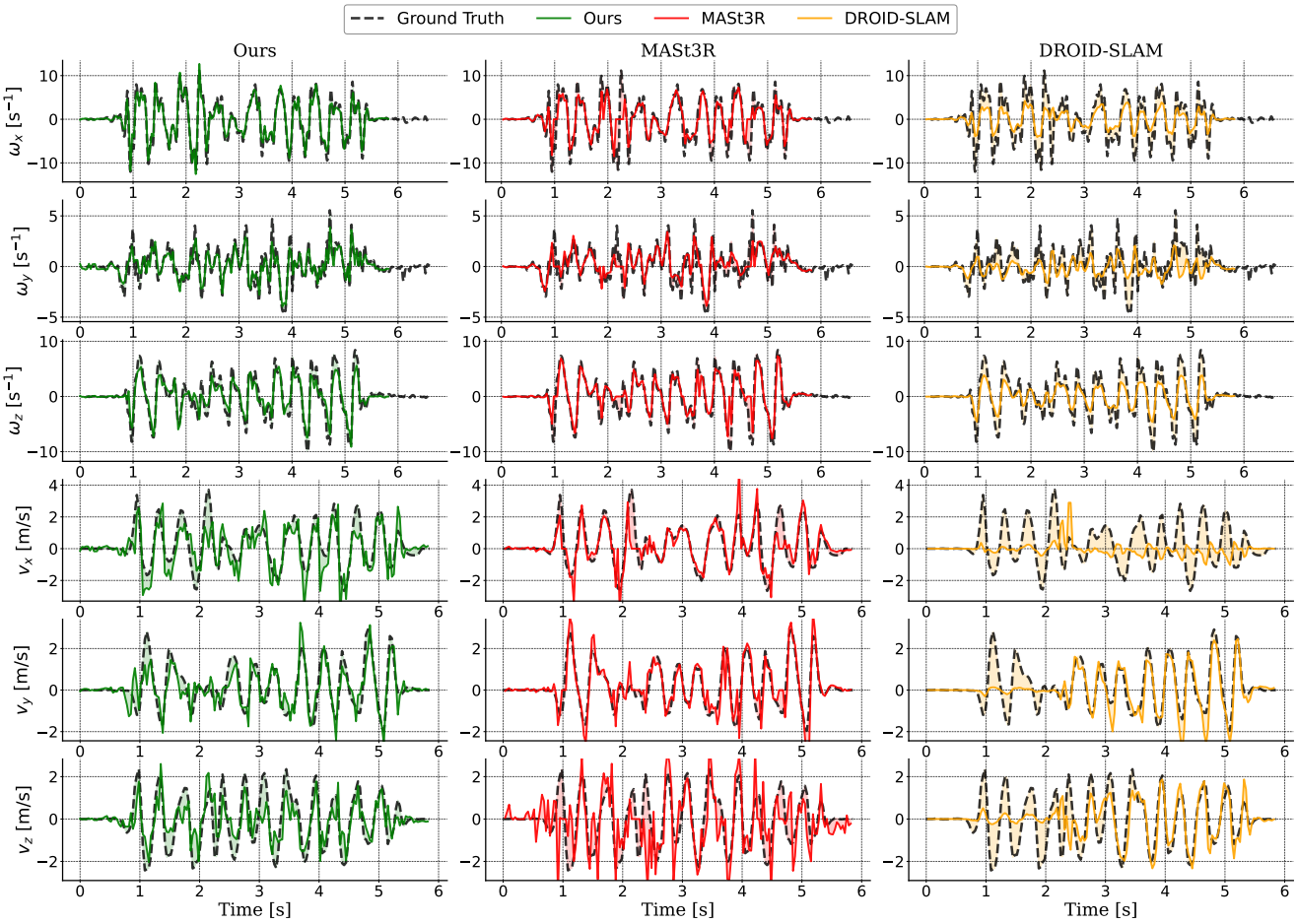

Figure 3. Velocity prediction comparison for the dining sequence.

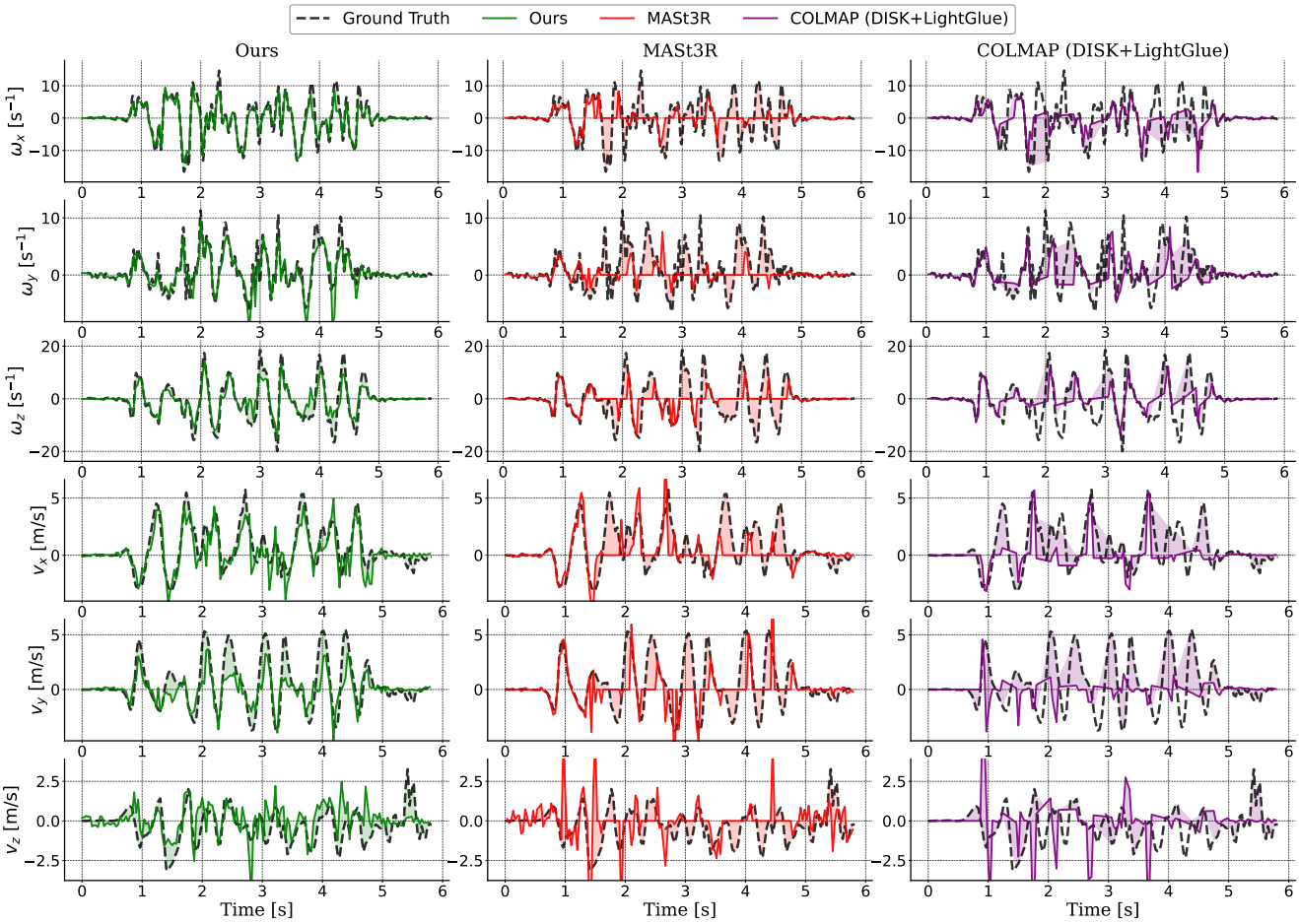

Figure 4. Velocity prediction comparison for the office sequence.