IMU-Assisted Learning of Single-View Rolling Shutter Correction, Supplementary Material

Jiawei Mo, Md Jahidul Islam, Junaed Sattar
Department of Computer Science and Engineering
University of Minnesota, Twin Cities, United States of America
[moxxx066, islam034, junaed]@umn.edu

Individual Study of RsDepthNet and RsPoseNet

We study the proposed RsDepthNet and RsPoseNet individually. To test one network, we disable the other and use its ground-truth data directly for rolling shutter correction. Results are reported in Table 1.

Table 1: Rolling shutter correction using predicted/ground-truth pose/depth.

Seq.		Input	pred. pose pred depth	pred. pose gt depth	gt pose pred depth	gt pose gt depth
02	EPE	5.738	0.937	0.902	0.142	0
	Ratio	-	99.3%	99.3%	100%	100%
07	EPE	6.593	0.746	0.726	0.111	0
	Ratio	-	95.9%	95.9%	100%	100%

The results indicate that the accuracy of rolling shutter correction by the proposed method is mainly limited by pose prediction using RsPoseNet and it is not very sensitive to depth accuracy.

If we expand Eq. 3 in the main paper, we get the equation that maps the input RS pixel $\mathbf{u}$ to the corrected GS pixel $\mathbf{u}'$:

$$d'\mathbf{u}' = \mathbf{K}\mathbf{R}[d\mathbf{K}^{-1}\mathbf{u}] + \mathbf{K}\mathbf{t}, \tag{1}$$

$$\mathbf{u}' \sim \mathbf{K}\mathbf{R}\mathbf{K}^{-1}\mathbf{u} + \frac{1}{d}\mathbf{K}\mathbf{t}. \tag{2}$$

The average linear velocity of Seq. 02 and Seq. 07 is $1.4m/s$ and $1.1m/s$, respectively. The readout time of the entire image is about 0.03 seconds. Therefore, the $\mathbf{t}$ in Eq. 2 is smaller than $0.05m$ and $\mathbf{K}\mathbf{t}$ is smaller than 12 pixels. The average depth d of Seq. 02 and Seq. 07 is $1.8m$ and $2.0m$, respectively. If we let $d = 1.9m$, any depth prediction between $1.643m$ to $2.257m$ results in sub-pixel error on $\mathbf{u}'$. Consequently, rolling shutter correction for our dataset is not very sensitive to depth prediction.

5th Conference on Robot Learning (CoRL 2021), London, UK.