

Supplementary Materials: Event-Guided Rolling Shutter Correction with Time-Aware Cross-Attentions

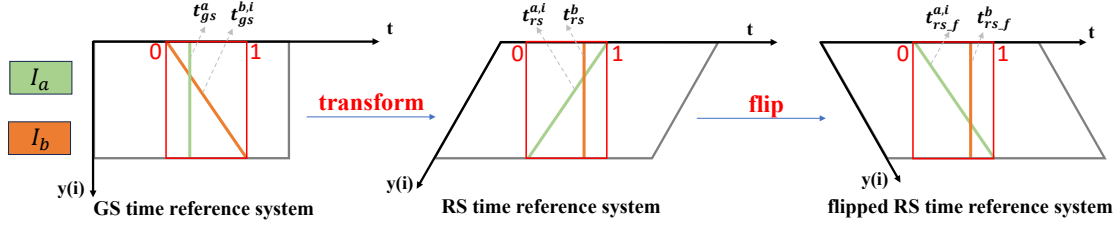


Figure 1: The process of self-supervised methods.

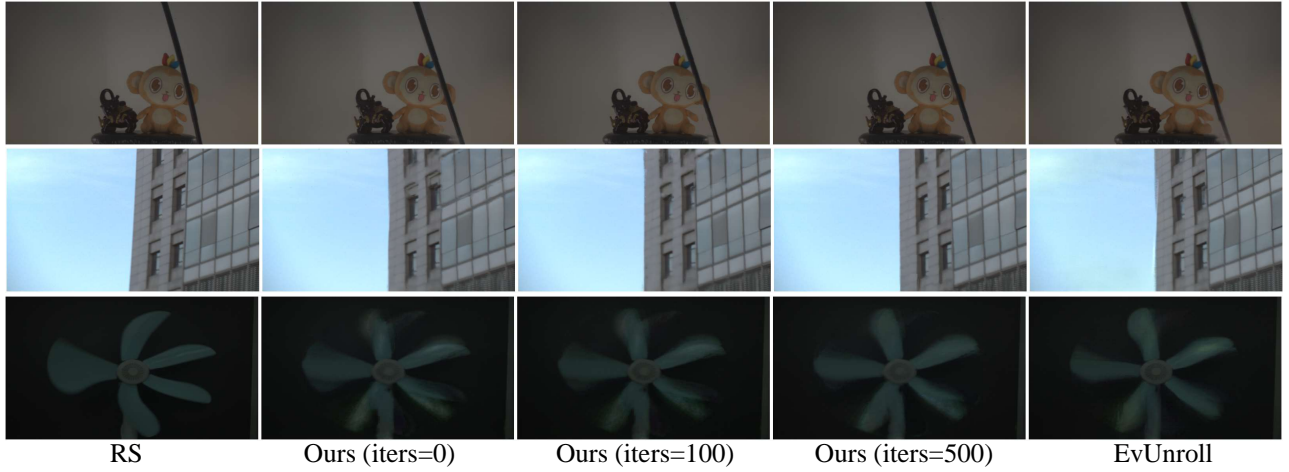


Figure 2: Qualitative comparison on real-world data.

1 IMPLEMENTATION DETAILS OF SELF-SUPERVISED METHODS

As shown in Fig. 1, in the GS time reference system, each row in GS image I_a has the same exposure time while each row in RS image I_b has different time as shown below. By normalizing the time dimension, we have $t_{gs}^{b,i} = \frac{i}{H-1}$, $i = (0, \dots, H-1)$.

After transformed to RS time reference system, I_b becomes "GS" with same exposure time for each row, and I_a becomes "RS", $t_{rs}^{a,i} = 1 - \frac{i}{H-1}$, $i = (0, \dots, H-1)$. To maintain the consistency of the RS time information in both system and reduce the training burden, we further flip the y-axis of the image and event and have a flipped RS time reference system. After the flip, $t_{rs_f}^{a,i} = t_{gs}^{b,i} = \frac{i}{H-1}$, and the GS time information changes, $t_{rs_f}^b = 1 - t_{gs}^a$.

After the flip operation, Eq. 11 and Eq. 12 were simplified, and can be rewritten as follows:

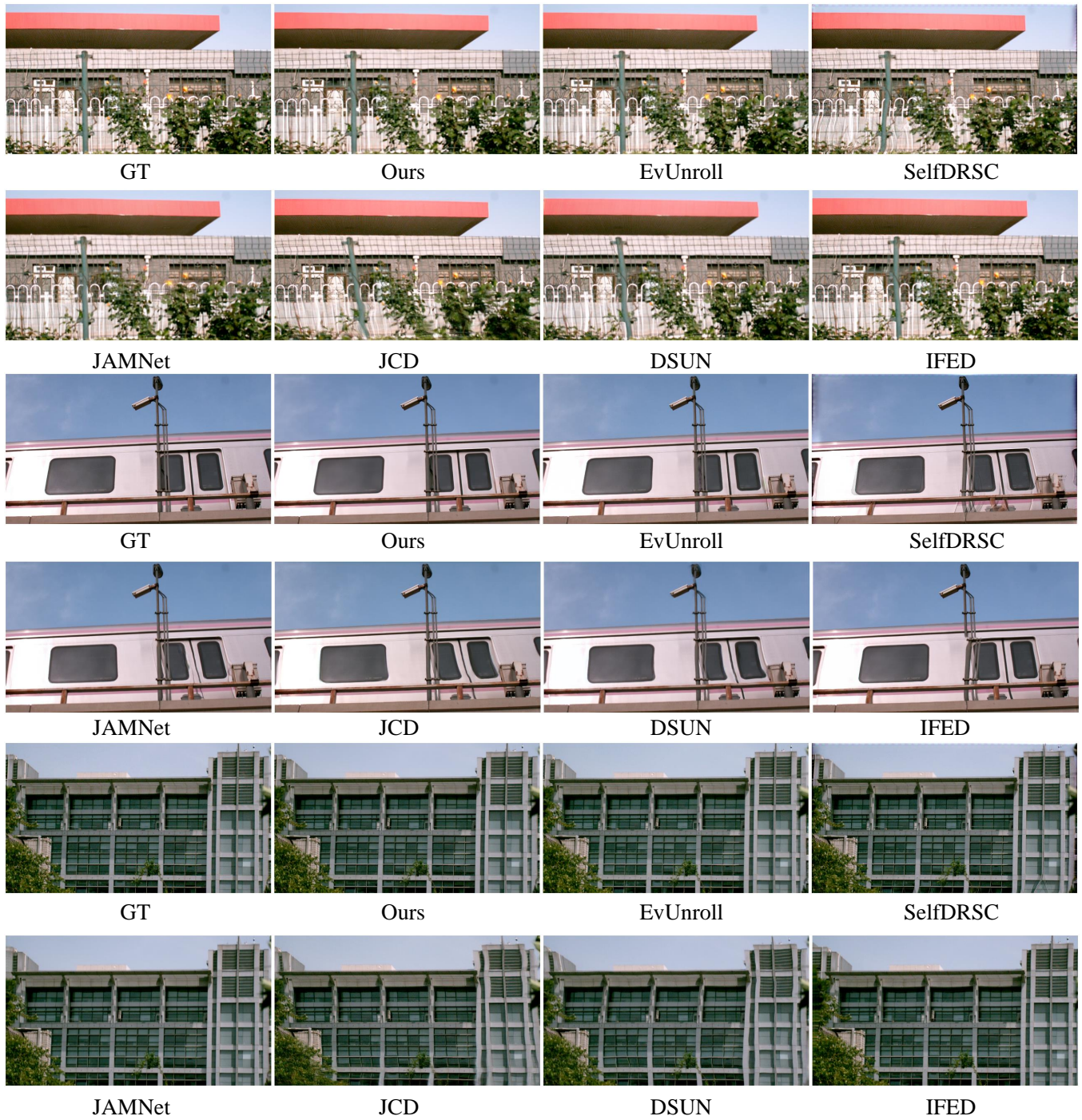
$$I_{o1} = f_{RSC}(I_a, e, t_{gs}, t_{rs}), I_{o2} = f_{RSC}(flip(I_{o1}), flip(e'), 1 - t_{gs}, t_{rs})$$

where $t_{rs} = \frac{i}{H-1}$ represents the timestamp of input RS image, and t_{gs} represents the timestamp of output GS image.

2 DISCUSSION ON SELF-SUPERVISION

As shown in Fig. 2, we present results from unsupervised training at different numbers of iteration. It is evident that even with just a few hundred iterations of unsupervised training, significant improvements can be achieved in removing artifacts and improving image quality.

SelfDRSC [1] also introduced a self-supervised training strategy. They adopt two dual RS images as input, as mentioned in Section.2 in our paper. In their method, the resulting GS output are warped back to RS using a bidirectional warping module that only employed during training, and the warped RS and input RS are used for self-supervised loss function based on temporal consistency, which shares certain similarity with our method in cyclic supervision. However, their method requires additional optical flow estimation module to conduct bidirectional warping for imposing self-supervision, whose accuracy and robustness would be critical to their RSC performance and multiple warping steps would cause error accumulation. In contrast, the self-supervision in our approach is achieved by transforming time reference system and does not involve any additional module. During training, our method involves just one additional forward pass. Some examples are illustrated in Fig. 3.



REFERENCES

- [1] Wei Shang, Dongwei Ren, Chaoyu Feng, Xiaotao Wang, Lei Lei, and Wangmeng Zuo. 2023. Self-supervised Learning to Bring Dual Reversed Rolling Shutter Images Alive. In *IEEE/CVF International Conference on Computer Vision, ICCV 2023, Paris*.

France, October 1-6, 2023. IEEE, 13040–13048. <https://doi.org/10.1109/ICCV51070.2023.01203>