Bringing Rolling Shutter Images Alive with Dual Reversed Distortion

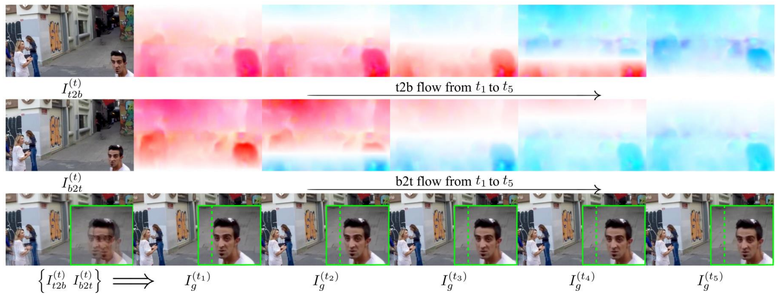
Zhihang Zhong, Mingdeng Cao, Xiao Sun, Zhirong Wu, Zhongyi Zhou, Yinqiang Zheng, Stephen Lin, and Imari Sato

GitHub: https://github.com/zzh-tech/Dual-Reversed-RS

The University of Tokyo (UTokyo), National Institute of Informatics (NII), Microsoft Research Asia (MSRA)

Extracting video clips from RS distorted images

- Rolling shutter (RS) distortion can be interpreted as the result of picking a row of pixels from instant global shutter (GS) frames over time during the exposure of the RS camera. This means that the information of each instant GS frame is partially, yet sequentially, embedded into the row-dependent distortion. Inspired by this fact, we address the challenging task of reversing this process, i.e., extracting undistorted GS frames from images suffering from RS distortion.



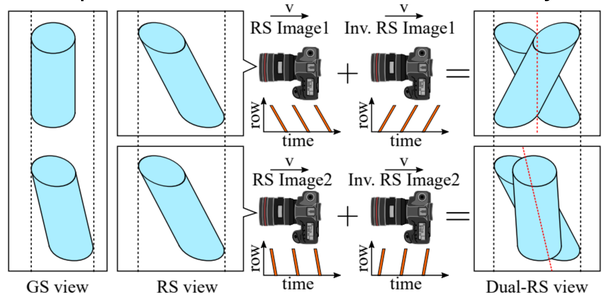
- However, since RS distortion is coupled with other factors such as readout settings and relative velocity, models that only exploit the geometric correlation between temporally adjacent images suffer from poor generality in processing data with different readout settings and dynamic scenes.

Summarized contributions

- **This is the first work that can extract video clips from distorted image in dynamic scenes. Besides, our solution can overcome the generalization problem caused by distinct readout settings**
- **We propose a novel end-to-end network architecture (IFED) that can iteratively estimate the accurate dual optical flow cube using pre-defined time cube prior for latent GS frame extraction.**
- **Extensive experimental results demonstrate the superior accuracy and robustness of IFED against the state-of-the-art both on synthetic dataset and real-world data.**

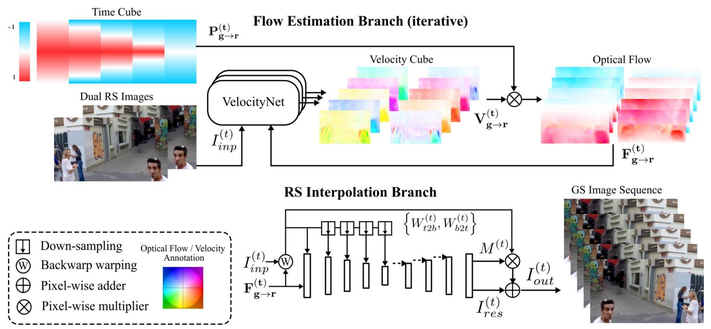
Merit of dual-RS to overcome correction ambiguity

- Correction ambiguity of consecutive frames setup: Suppose there are two similar cylinders, one of them is tilted, as shown in GS view. Then, two RS cameras moving horizontally at the same speed but with different readout time setups can produce the same RS view. Models do not know how much correction is correct facing data beyond their training dataset.
- Instead of two consecutive frames, we introduce another constraint setting that utilizes intra-frame spatial constraints of dual images taken simultaneously but with reversed distortion. Dual-RS setup can avoid ambiguity because the correct correction pose can be estimated based on the symmetry



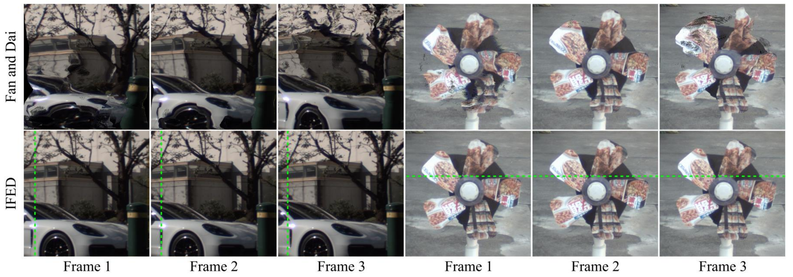
Network architecture (IFED)

- We propose a dual time cube as an RS prior, regress the dual velocity cube to indirectly estimate the dual optical flow for backward warping, and then through an encoder-decoder branch, efficiently merge the symmetric information to extract the potential GS frames.



Experimental results

- Previous method cannot generalize to either the case of camera-only motion (the left example) or the case of object-only motion (the right example), while ours is robust to different motion patterns.



- Grounded on the symmetric and complementary nature of dual reversed distortion, our method can successfully generalize to different readout settings without artifacts and undesired distortions.

Conclusions and future works

In this work, we addressed a challenging task of restoring consecutive distortion-free frames from RS distorted images in dynamic scenes. We designed an end-to-end deep neural network IFED for the dual-RS setup, which has the advantages of being able to model dynamic scenes and not being affected by distinct readout times. How to better model in the time dimension to achieve the recovery of GS frames at arbitrary exposure moments is our future work.