Context-Aware Video Reconstruction for Rolling Shutter Cameras

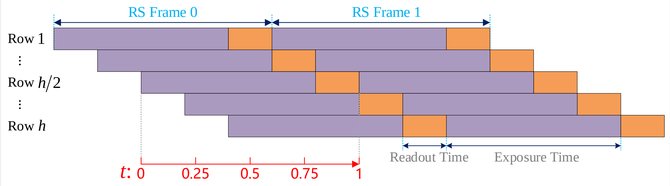
Bin Fan, Yuchao Dai*, Zhiyuan Zhang, Qi Liu, Mingyi He

School of Electronics and Information, Northwestern Polytechnical University, Xi'an, China

https://github.com/GitCVfb/CVR

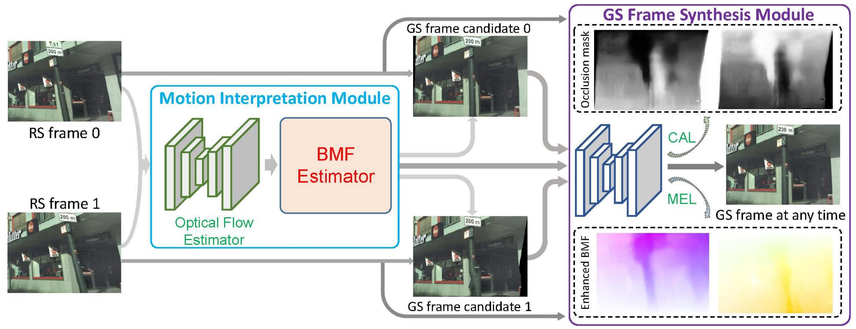
Problem Statement

Given two rolling shutter (RS) images at adjacent times 0 and 1, we aim to synthesize an intermediate global shutter (GS) frame corresponding to any time t, where $0 \le t \le 1$.



Note that the middle scanlines of the two RS frames correspond to time instances 0 and 1, respectively.

Network Framework



- ✓ Bidirectional optical flow estimator.
- ✓ Bilateral motion field estimator (NBMF or ABMF) at arbitrary time t∈[0,1].
- ✓ GS frame synthesis module, including a motion enhancement layer (MEL) and a contextual aggregation layer (CAL).

Bilateral Motion Field (BMF) Estimator

- ➢ The bilateral motion field can be generated by scaling the regular optical flow field, i.e.,

$$\mathbf{U}_{0\to t}(\mathbf{x}) = \mathbf{C}_{0\to t}(\mathbf{x}) \cdot \mathbf{F}_{0\to 1}(\mathbf{x})$$
$$\mathbf{U}_{1\to t}(\mathbf{x}) = \mathbf{C}_{1\to t}(\mathbf{x}) \cdot \mathbf{F}_{1\to 0}(\mathbf{x})$$

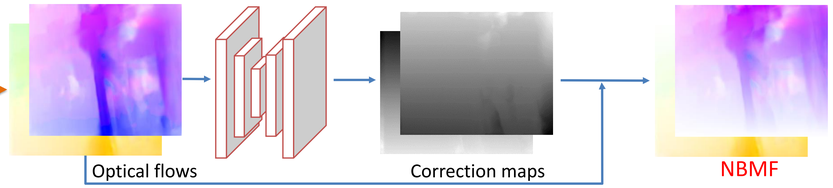
- ➢ The bilateral correction map was formulated under the constant camera motion in *ICCV2021*, i.e.,

$$\mathbf{C}_{0\to t}(\mathbf{x}) = \frac{(t-\tau_0)(h-\pi_v)}{h}$$
$$\mathbf{C}_{1\to t}(\mathbf{x}) = \frac{(\tau_1-t)(h+\pi'_v)}{h}$$

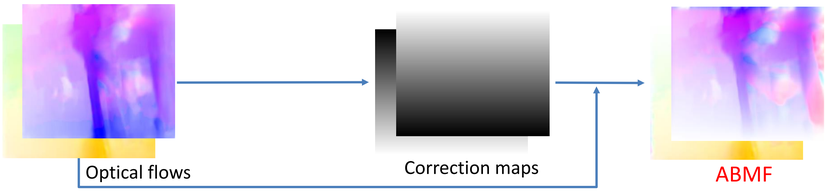
- ➢ In this paper, we further propose its approximated version by neglecting the parallax effects, i.e.,

$$\mathbf{C}_{0\to t}(\mathbf{x}) = t-\tau_0$$
$$\mathbf{C}_{1\to t}(\mathbf{x}) = \tau_1-t$$

- **Network-based BMF (NBMF)**



- **Approximated BMF (ABMF)**



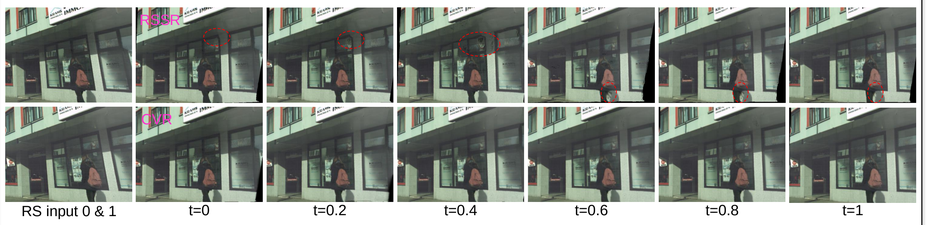
Experimental Results

Performance on Carla-RS and Fastec-RS datasets

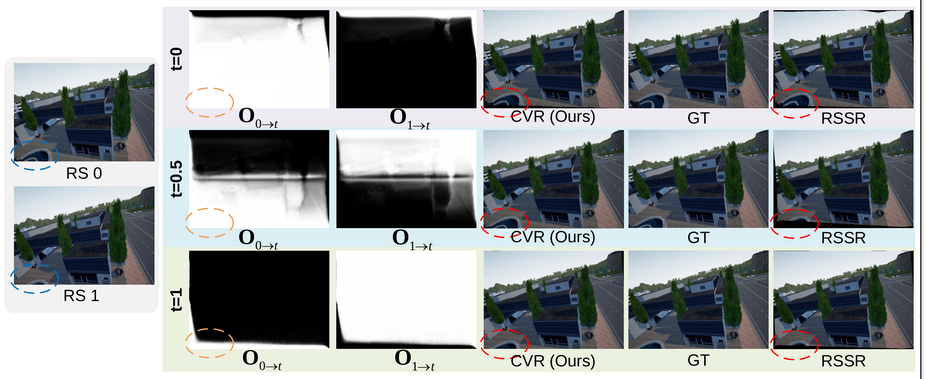
Method	Runtime (seconds)	PSNR↑ (dB)			SSIM↑		LPIPS↓	
		CRM	CR	FR	CR	FR	CR	FR
DiffSfM [61]	467	24.20	21.28	20.14	0.775	0.701	0.1322	0.1789
DiffHomo [62]	424	19.60	18.94	18.68	0.606	0.609	0.1798	0.2229
DeepUnrollNet [24]	0.34	26.90	26.46	26.52	0.807	0.792	0.0703	0.1222
SUNet [10]	0.21	29.28	29.18	28.34	0.850	0.837	0.0658	0.1205
RSSR*	0.09	28.20	23.86	21.02	0.839	0.768	0.0764	0.1866
RSSR [9]	0.12	30.17	24.78	21.23	0.867	0.776	0.0695	0.1659
*CVR** (Ours)	0.12	31.82	31.60	28.62	0.927	0.845	0.0372	0.1117
CVR (Ours)	0.14	**32.02**	**31.74**	**28.72**	**0.929**	**0.847**	**0.0368**	**0.1107**

*: *applying our proposed approximated bilateral motion field (ABMF) model.*

Generating high-fidelity global shutter video sequences



Intermediate process of occlusion reasoning $I_t^g = O_{0\to t} I_{0\to t}^g + O_{1\to t} I_{1\to t}^g$



Generalizability on real data