Deblur e-NeRF: NeRF from Motion-Blurred Events under High-speed or Low-light Conditions

Weng Fei Low   Gim Hee Lee

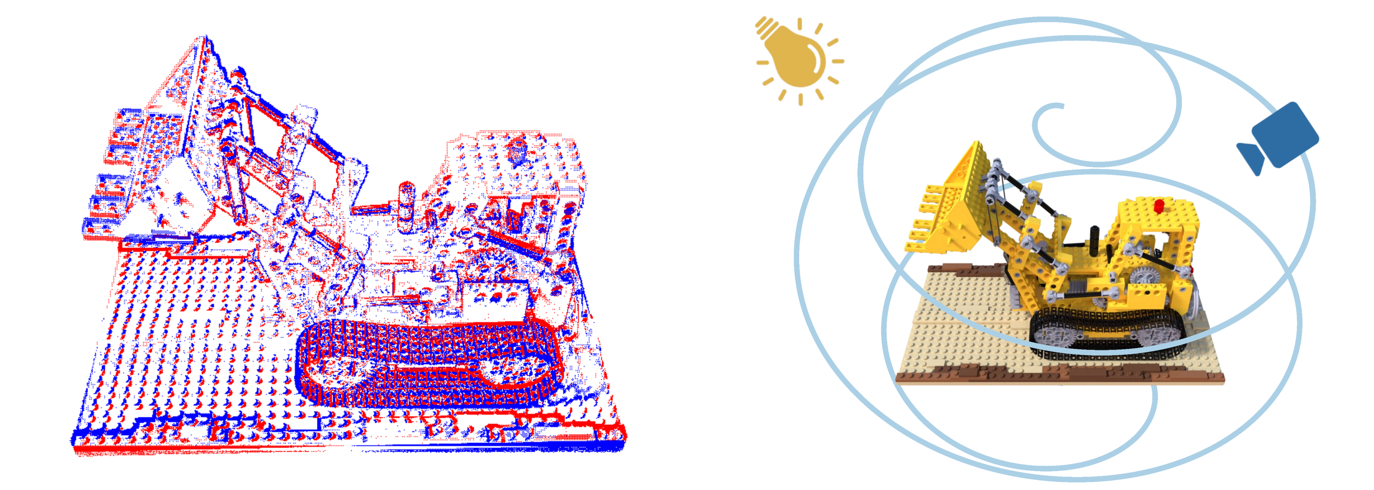
National University of Singapore

ECCV MILANO 2024

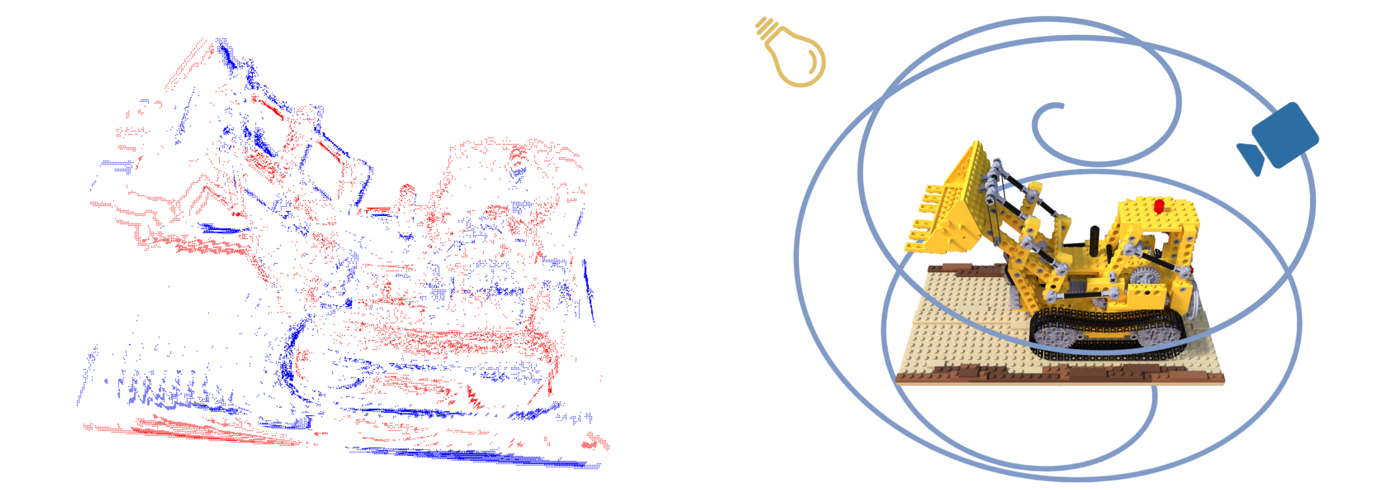
Motivation

1. Event cameras outperform standard cameras under:
   - High speed
   - High dynamic range
   - Low light

   due to their distinctive principle of operation
2. However, event cameras also suffer from motion blur, especially under these challenging conditions
3. This is due to the limited bandwidth of the event sensor pixel, which is mostly proportional to the light intensity



(a) Minimally motion-blurred events (left), under low speed and bright light (right).

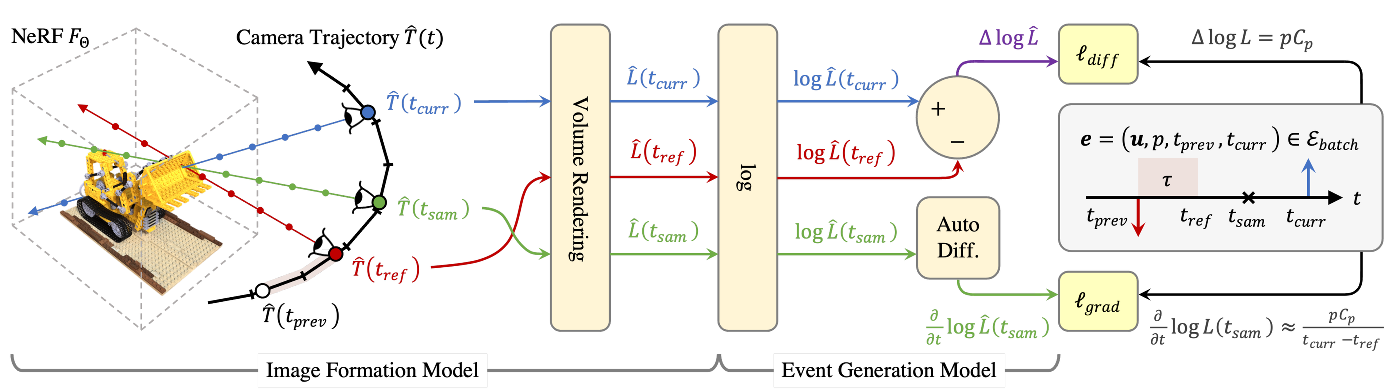


(b) Significantly motion-blurred events (left), under high speed and low light (right).

Research Question

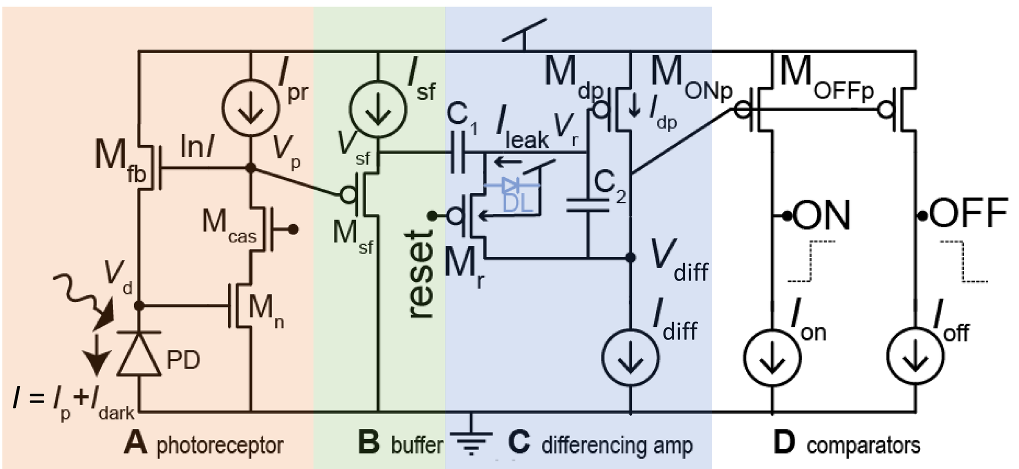
How to reconstruct blur-minimal NeRFs from motion-blurred events, generated under high-speed or low-light conditions?

Prior Work: Robust e-NeRF



Proposed Method: Deblur e-NeRF

Pixel Bandwidth Model



Input: $u = \log L$

State: $\boldsymbol{x} = \left[\partial \log L_p / \partial t \;\; \log L_p \;\; \log L_{sf} \;\; \log L_{diff} \right]^\top$

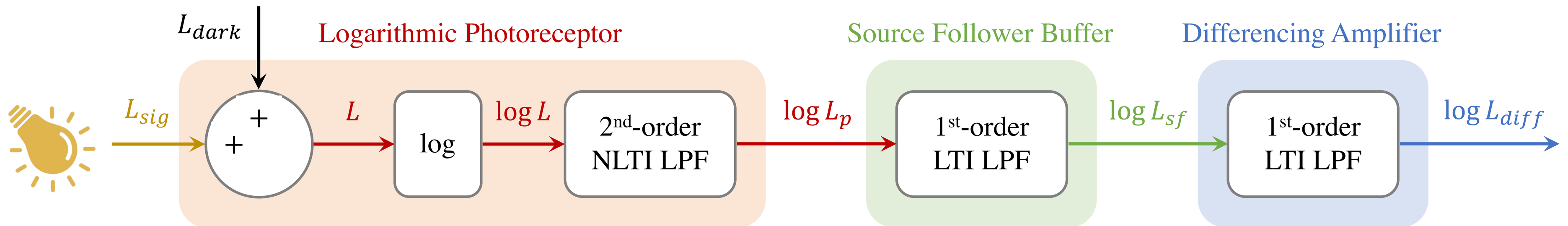
Output: $\boldsymbol{y} = \left[\log L_{sf} \;\; \log L_{diff} \right]^\top$

State-Space Model:

$$\dot{\boldsymbol{x}}(t) = A\left(u(t)\right) \boldsymbol{x}(t) + B\left(u(t)\right) u(t)$$

$$\boldsymbol{y}(t) = C \, \boldsymbol{x}(t) \text{ , where:}$$

$$A(u) = \begin{bmatrix} -2\zeta(u)\omega_n(u) & -\omega_n^2(u) & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & \omega_{c,sf} & -\omega_{c,sf} & 0 \\ 0 & 0 & \omega_{c,diff} & -\omega_{c,diff} \end{bmatrix}, \quad B(u) = \begin{bmatrix} \omega_n^2(u) \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

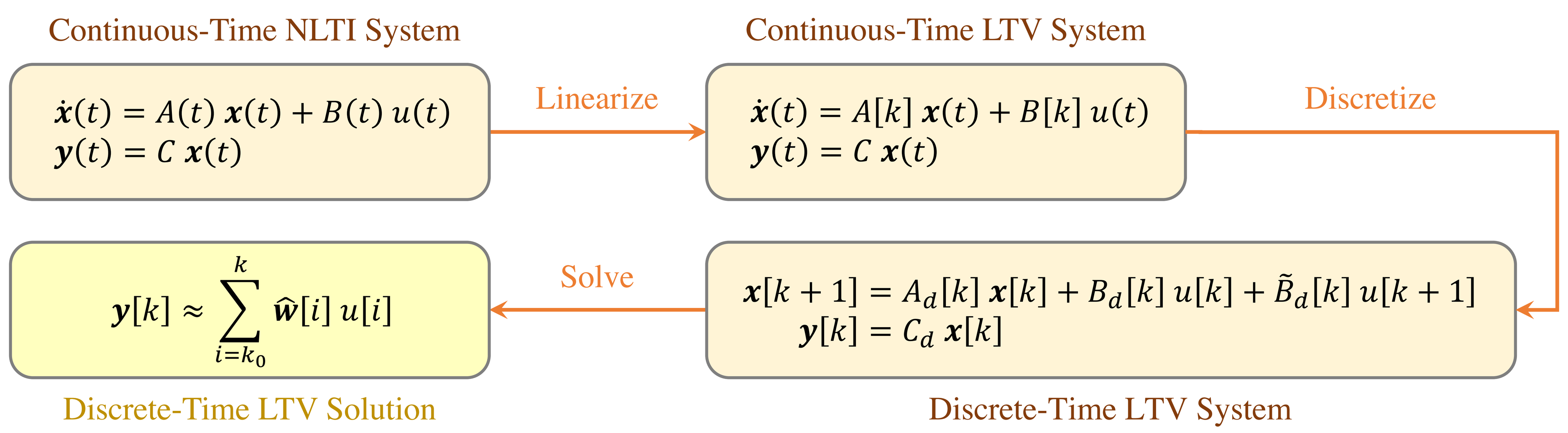


Steady-State Model of Differencing Amplifier Reset Mechanism:

$$\log L_{blur}(t) = \log L_{diff}(t) + \log L_{delta}(t_{ref}) \; e^{-\omega_{c,diff}(t - t_{ref})}$$

where $\log L_{delta} = \log L_{sf} - \log L_{diff}$ and $t \geq t_{ref}$.

Synthesis of Motion-Blurred Effective Log-Radiance



Importance Sampling: $T_i \sim \mathrm{Exp}\left(t_k - t_i;\; \omega_{c,dom,min}\right) = \omega_{c,dom,min} \; e^{-\omega_{c,dom,min}(t_k - t_i)}$

Threshold-Normalized Total Variation Loss

$$\ell_{tv}(\boldsymbol{e}) = \left| \frac{\delta \log \hat{L}_{blur}}{\bar{C}} \right| \text{ , where } \bar{C} = \tfrac{1}{2}(C_{-1} + C_{+1}),$$

$$\delta \log \hat{L}_{blur} := \log \hat{L}_{blur}(\boldsymbol{u}, t_{end}) - \log \hat{L}_{blur}(\boldsymbol{u}, t_{start}).$$

Translated-Gamma Correction

$$\hat{\boldsymbol{L}}_{sig,corr} = \boldsymbol{b} \odot \hat{\boldsymbol{L}}^{a} - \boldsymbol{c} \text{ , where}$$

a, $\boldsymbol{b}$ and $\boldsymbol{c}$ are correction parameters.

Novel View Synthesis Results

Table 1: Upper bound performance without event motion blur

Method	PSNR ↑	SSIM ↑	LPIPS ↓
E2VID + NeRF	19.49	0.847	0.268
Robust e-NeRF	28.48	0.944	0.054
Deblur e-NeRF	**29.43**	**0.953**	**0.043**

Table 2: Quantitative results of the real exps.

Method	08_peanuts_running			11_all_characters		
	PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓
E2VID + NeRF	14.85	0.690	0.595	13.12	0.695	0.627
Robust e-NeRF	18.00	0.677	0.507	15.91	0.677	0.552
Deblur e-NeRF	**18.27**	**0.695**	**0.503**	**16.53**	**0.710**	**0.511**

Table 3: Effect of camera speed. †Trained with 1/8× the batch size of baselines.

Method	$v = 0.125\times$			$v = 1\times$			$v = 4\times$		
	PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓
E2VID + NeRF	18.58	0.849	0.259	18.85	0.839	0.278	17.82	0.804	0.328
Robust e-NeRF	28.31	0.943	0.050	26.11	0.924	0.074	22.18	0.861	0.122
Deblur e-NeRF†	**28.71**	**0.948**	**0.048**	**28.41**	**0.947**	**0.049**	**27.48**	**0.939**	**0.061**

Table 4: Effect of scene illuminance. †Trained with 1/8× the batch size of baselines.

Method	$E_{sc} = 100\ 000 lux$			$E_{sc} = 1\ 000 lux$			$E_{sc} = 10 lux$		
	PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓
E2VID + NeRF	19.27	0.846	0.268	18.85	0.839	0.278	17.24	0.804	0.354
Robust e-NeRF	27.62	0.942	0.055	26.11	0.924	0.074	22.72	0.870	0.129
Deblur e-NeRF†	**28.73**	**0.948**	**0.047**	**28.41**	**0.947**	**0.049**	**28.62**	**0.935**	**0.059**

Table 5: Collective effect of camera speed and scene illuminance. †Trained with 1/8× the batch size of baselines.

Method	Opt. C_p & τ	Opt. Ω	$v = 0.125\times$, $E_{sc} = 100\ 000 lux$			$v = 1\times$, $E_{sc} = 1\ 000 lux$			$v = 4\times$, $E_{sc} = 10 lux$		
			PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓
E2VID + NeRF	–	–	19.19	0.844	0.281	18.85	0.839	0.278	15.37	0.799	0.436
Robust e-NeRF	×	–	28.27	0.944	0.057	26.11	0.924	0.074	18.42	0.814	0.255
	✓	–	28.28	0.944	0.051	26.31	0.923	0.075	18.51	0.812	0.254
Deblur e-NeRF†	×	×	**29.00**	**0.950**	**0.043**	**28.41**	**0.947**	**0.049**	**26.15**	**0.904**	**0.134**
	×	✓	28.19	0.943	0.046	26.07	0.930	0.067	25.59	0.896	0.156