

Weng Fei Low Gim Hee Lee

Motivation

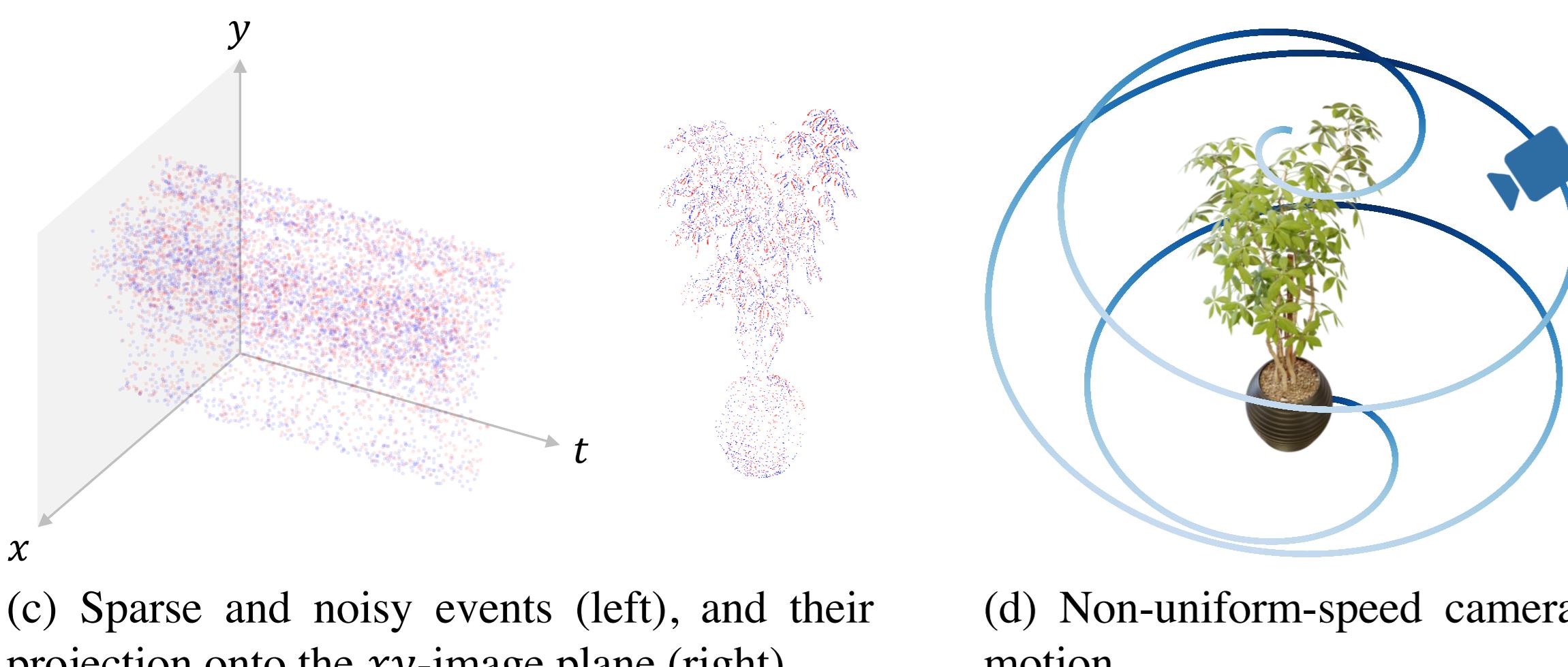
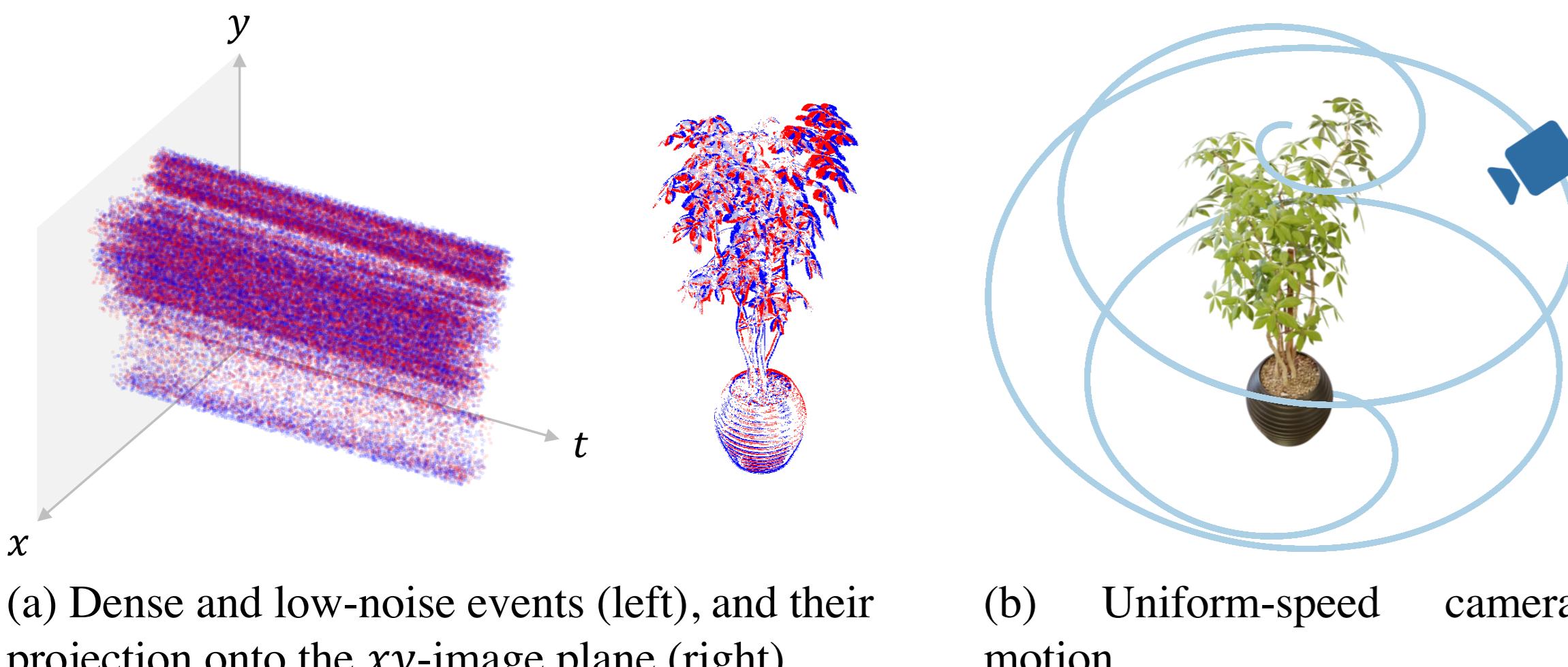
- Event cameras offer many advantages over standard cameras due to their distinctive principle of operation:
 - Low power
 - High temporal resolution
 - Low latency
 - High dynamic range
- Many downstream visual applications hinges on an efficient & effective scene representation, where Neural Radiance Field (NeRF) is seen as the leading candidate

Research Question

How to robustly reconstruct a NeRF from a moving event camera under general real-world conditions?

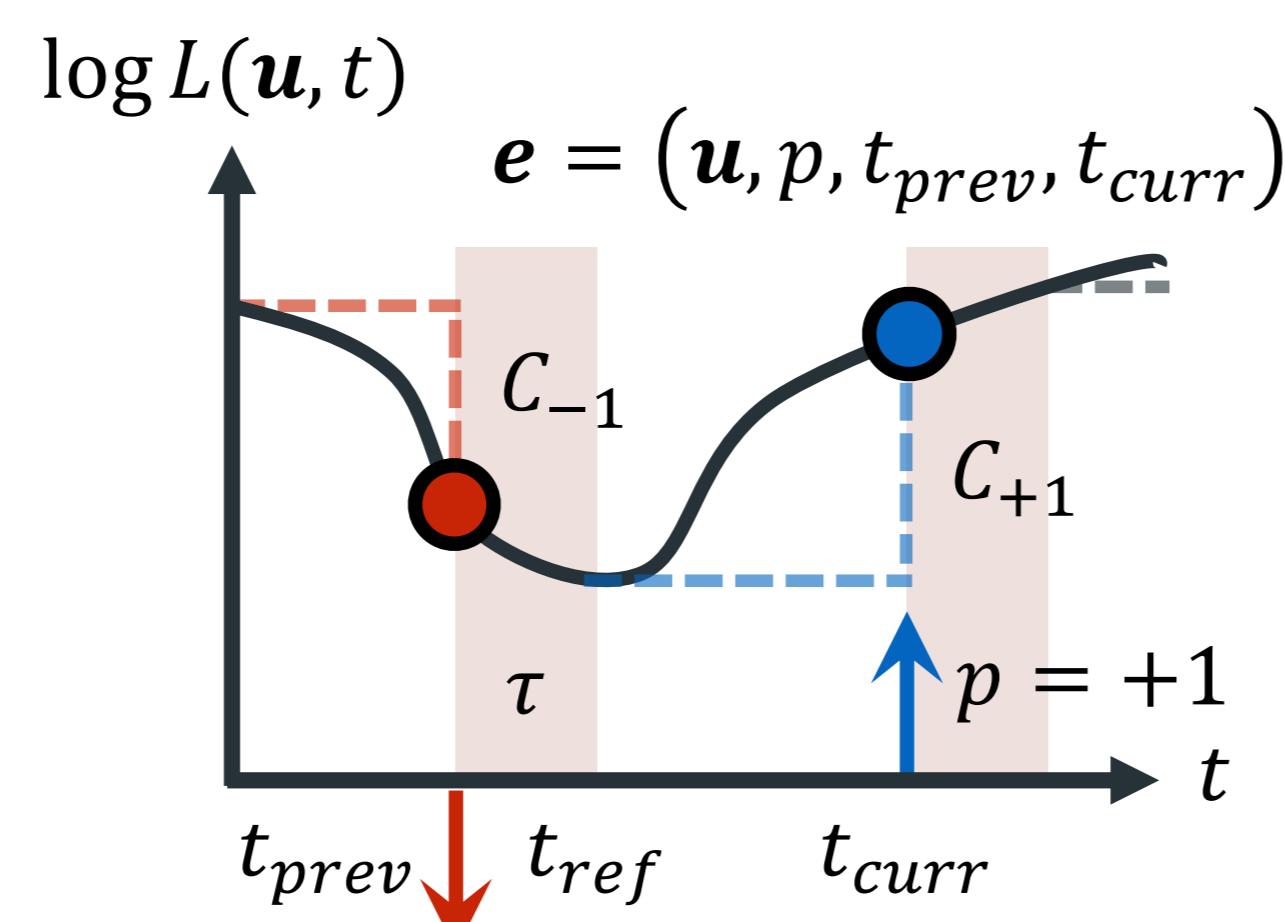
Limitations of Prior Work

- Depends on a temporally dense & low-noise event stream
- Does not directly & effectively generalize to arbitrary contrast threshold values & camera speed profiles



Proposed Method: Robust e-NeRF

Event Generation Model

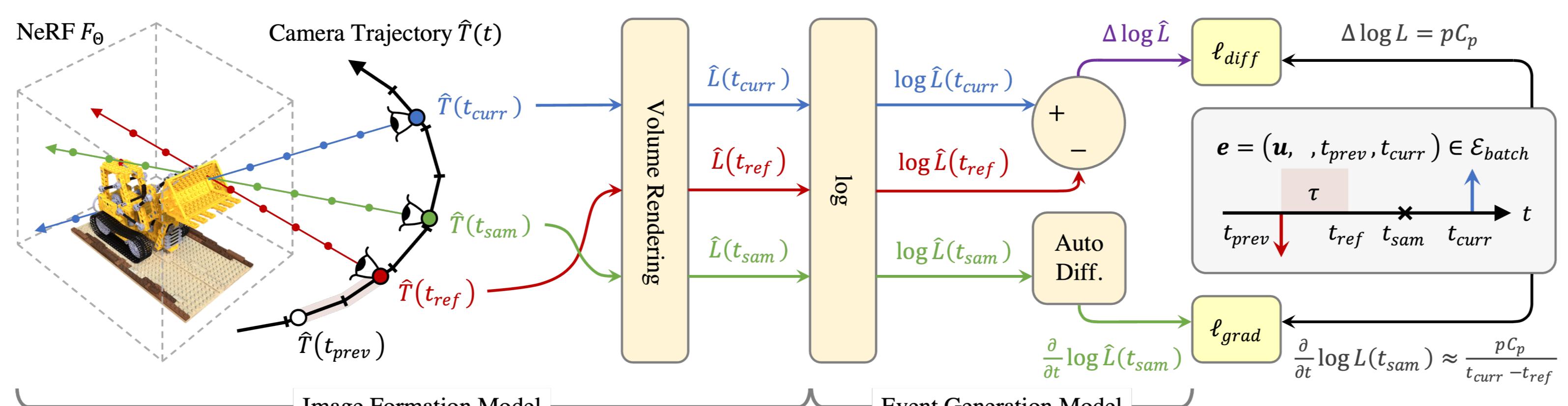


- e is an Event generated by pixel \mathbf{u} , with polarity $p \in \{-1, +1\}$, at timestamp t_{curr}
- Measured log-radiance difference: $\Delta \log L := \log L(\mathbf{u}, t_{curr}) - \log L(\mathbf{u}, t_{ref}) = pC_p$
- Reference log-radiance timestamp: $t_{ref} = t_{prev} + \tau$

✓ Accounts for: t_{prev} is the previous event timestamp at \mathbf{u}

- Intrinsic parameters:** Time-independent, asymmetric Contrast Threshold C_p
Refractory Period τ (i.e. pixel dead-time \rightarrow sparsity)
- Non-idealities:** Pixel-to-pixel contrast threshold variation σ_{C_p} (i.e. noise)

Training Pipeline



$$\text{Threshold-Normalized Difference Loss: } \ell_{diff}(e) = \left(\frac{\Delta \log \hat{L} - pC_p}{\bar{C}} \right)^2, \bar{C} = \frac{1}{2}(C_{-1} + C_{+1})$$

$$\text{Target-Normalized Gradient Loss: } \ell_{grad}(e) = \text{APE} \left(\frac{\partial}{\partial t} \log \hat{L}(\mathbf{u}, t_{sam}), \frac{pC_p}{t_{curr} - t_{ref}} \right) \text{ where } \text{APE}(\hat{y}, y) = \left| \frac{\hat{y} - y}{y} \right|$$

- Effectively generalizes to arbitrary speed profiles & intrinsic parameter values, without such prior knowledge
- Does not involve accumulation of successive events, thereby reconstructions are detailed and robust to event sparsity & noise

Gamma Correction of Synthesized Views

$$\text{Affine-corrected log-radiance prediction: } \log \hat{L}_{corr} = \mathbf{a} \odot \log \hat{L} + \mathbf{b}$$

$$\text{Gamma-corrected linear-radiance prediction: } \hat{L}_{corr} = (\exp \mathbf{b}) \hat{L}^{\circ \mathbf{a}}$$

Novel View Synthesis Results

Method	$v = 1 \times$			$v_b = 8 \times$			$v = \frac{1}{8} \times$			$v = 8 \times$		
	PSNR \uparrow	SSIM \uparrow	LPIPS \downarrow	PSNR \uparrow	SSIM \uparrow	LPIPS \downarrow	PSNR \uparrow	SSIM \uparrow	LPIPS \downarrow	PSNR \uparrow	SSIM \uparrow	LPIPS \downarrow
E2VID + NeRF	18.92	0.832	0.316	18.92	0.832	0.316	18.92	0.832	0.316	18.92	0.832	0.316
Ev-NeRF	27.72	0.935	0.087	26.25	0.926	0.102	19.79	0.792	0.326	20.83	0.862	0.198
Robust e-NeRF	28.19	0.945	0.057	28.19	0.945	0.057	28.19	0.945	0.057	28.19	0.945	0.057

Method	Opt. C_p	$\sigma_{C_p} = 0.00$			$\sigma_{C_p} = 0.03$			$\sigma_{C_p} = 0.06$		
		PSNR \uparrow	SSIM \uparrow	LPIPS \downarrow	PSNR \uparrow	SSIM \uparrow	LPIPS \downarrow	PSNR \uparrow	SSIM \uparrow	LPIPS \downarrow
E2VID + NeRF	—	18.92	0.832	0.316	18.68	0.827	0.330	18.03	0.808	0.363
Ev-NeRF	✗	27.72	0.935	0.087	24.42	0.895	0.155	8.07	0.841	0.260
Robust e-NeRF	✗	27.43	0.911	0.123	23.66	0.826	0.261	15.43	0.708	0.441
	✗	28.19	0.945	0.057	28.14	0.946	0.058	28.23	0.947	0.057
	✓	28.17	0.946	0.051	27.91	0.946	0.054	28.19	0.948	0.049

Method	Opt. C_p	τ	$\tau = 0ms$			$\tau = 8ms$			$\tau = 25ms$		
			PSNR \uparrow	SSIM \uparrow	LPIPS \downarrow	PSNR \uparrow	SSIM \uparrow	LPIPS \downarrow	PSNR \uparrow	SSIM \uparrow	LPIPS \downarrow
E2VID + NeRF	—	—	18.92	0.832	0.316	14.87	0.797	0.427	14.15	0.791	0.467
Ev-NeRF	✗	—	27.72	0.935	0.087	13.17	0.707	0.559	12.75	0.759	0.528
Robust e-NeRF	✗	—	27.43	0.911	0.123	13.56	0.716	0.528	13.75	0.717	0.569
	✗	✗	28.19	0.945	0.057	26.30	0.934	0.066	25.51	0.929	0.072
	✓	✗	28.18	0.945	0.052	23.43	0.910	0.090	22.48	0.895	0.105

Method	Opt. C_p	τ	$v_b = 1 \times, \sigma_{C_p} = 0.00, \tau = 0ms$			$v_b = 4 \times, \sigma_{C_p} = 0.03, \tau = 8ms$			$v_b = 8 \times, \sigma_{C_p} = 0.06, \tau = 25ms$		
			PSNR \uparrow	SSIM \uparrow	LPIPS \downarrow	PSNR \uparrow	SSIM \uparrow	LPIPS \downarrow	PSNR \uparrow	SSIM \uparrow	LPIPS \downarrow
E2VID + NeRF	—	—	18.92	0.832	0.316	14.98	0.796	0.433	14.07	0.801	0.448
Ev-NeRF	✗	—	27.72	0.935	0.087	12.33	0.742	0.521	12.05	0.807	0.425
Robust e-NeRF	✗	—	27.43	0.911	0.123	13.06	0.732	0.539	12.27	0.772	0.539
	✗	✗	28.19	0.945	0.057	24.10	0.913	0.086	23.51	0.900	0.110