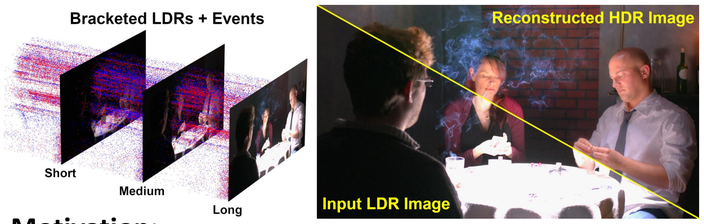
HDR Reconstruction from Bracketed Exposures and Events

Richard Shaw, Sibi Catley-Chandar, Aleš Leonardis, Eduardo Pérez-Pellitero

BMVC 2022

HDR from Images and Events



Motivation:

- Image-based HDR methods suffer from limited dynamic range per exposure bracket
- Motion related artefacts (ghosting)

Event-based Cameras:

- Capture motion in the scene, no motion → no signal
- Asynchronous independent pixels
- Event $E = \{x, y, t, p\}$ if brightness change > threshold

Advantages:

- High temporal resolution (~1μs)
- High dynamic range (140 dB vs 60 dB)
- Ultra-low power (1mW vs 1W)
- No motion blur

Limitations:

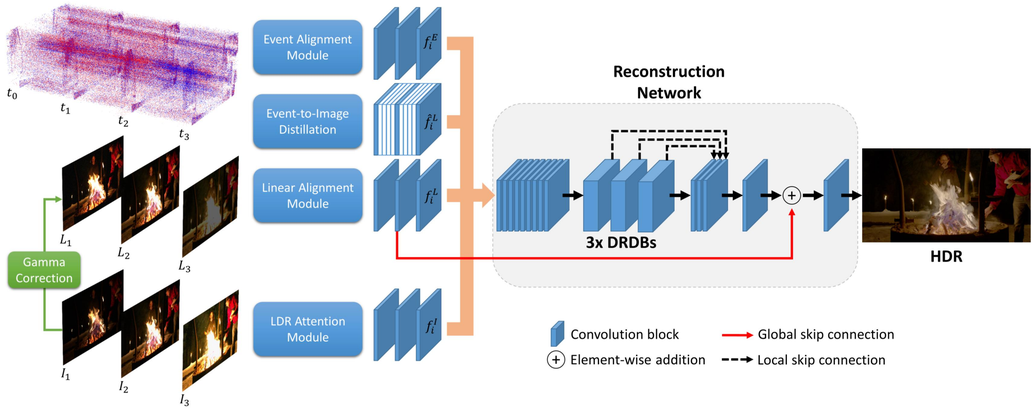
- Information only in moving regions
- Grayscale
- Low-resolution
- Fine detail difficult to reconstruct (SoTA results are poor quality)

Idea: Combine the best of both worlds

- A multi-modal HDR model that leverages both bracketed exposures from a conventional frame-based camera and events from an event-based camera
- Event cameras provide **high-frequency visual information**, images provide **rich RGB colour and fine details** at lower frequency

Our Method

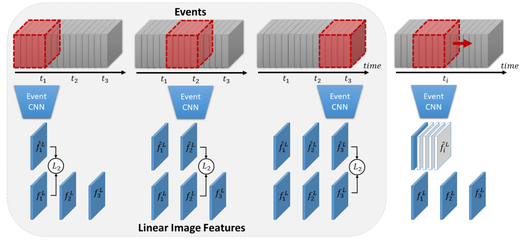
$$\hat{H} = g\left(\mathcal{A}(I_i), \mathcal{P}^L(L_i), \mathcal{P}^E(E_i), \mathcal{D}(E_j); \theta\right)$$



Event-to-Image Distillation:

- Leverage complementary information between events and images
- Translate sub-sampled event features into image features
- Self-supervising loss on the event features and corresponding image features:

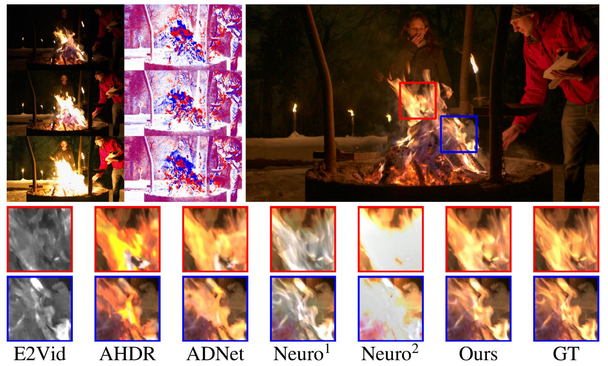
$$\mathcal{L}_{\mathcal{D}} = \sum_{s=1}^{S} \sum_{i=1}^{3} \left(f_{i,s}^{E} - \mathrm{sg}\left(f_{i,s}^{L}\right)\right)^2$$

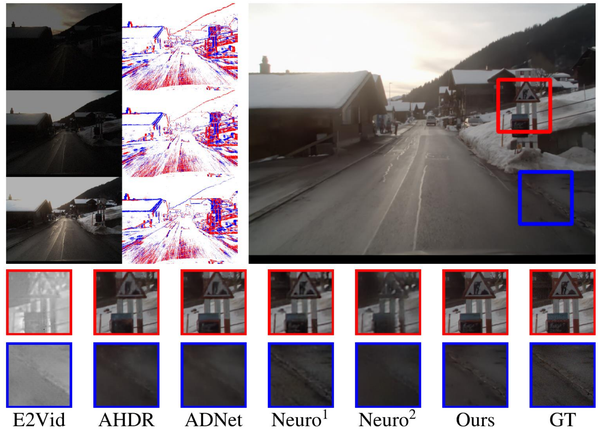


Ablations

Method	Imbalanced Params				Balanced Params			
	Params(M)	PSNR-L	PSNR-μ	HV2	Params(M)	PSNR-L	PSNR-μ	HV2
Images-only	2.81	39.18	36.89	50.06	6.14	37.75	36.33	48.92
+ Event alignment	4.38	40.97	37.35	55.14	6.27	39.55	36.62	52.15
+ Event sub-sampling	4.67	41.32	37.60	54.85	6.74	40.39	36.98	53.33
+ Event-to-image distill.	6.13	**41.81**	**37.84**	**55.79**	6.13	**41.81**	**37.84**	**55.79**

Results





Conclusions

- **Multi-model** HDR architecture combining events and low dynamic range images
- Leverages **high-frequency events** to better align and reconstruct details
- Significant improvements over other SoTA event- or image-only methods