Distil-E2D: Distilling Image-to-Depth Priors for Event-Based Monocular Depth Estimation

Jie Long Lee, Gim Hee Lee

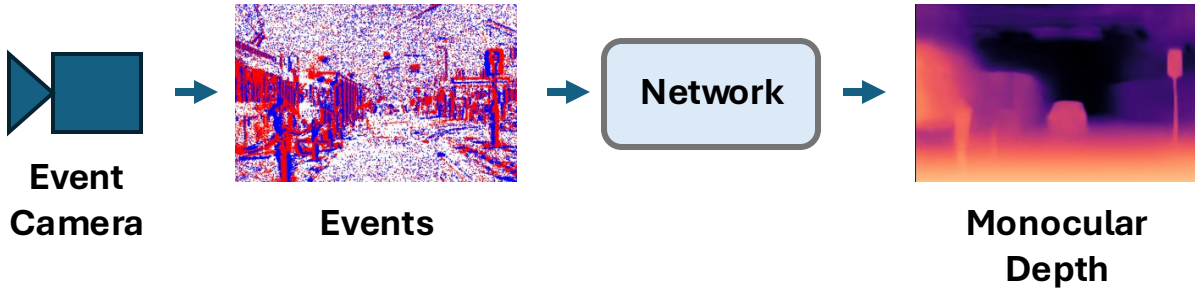
NeurIPS 2025

Problem Setting

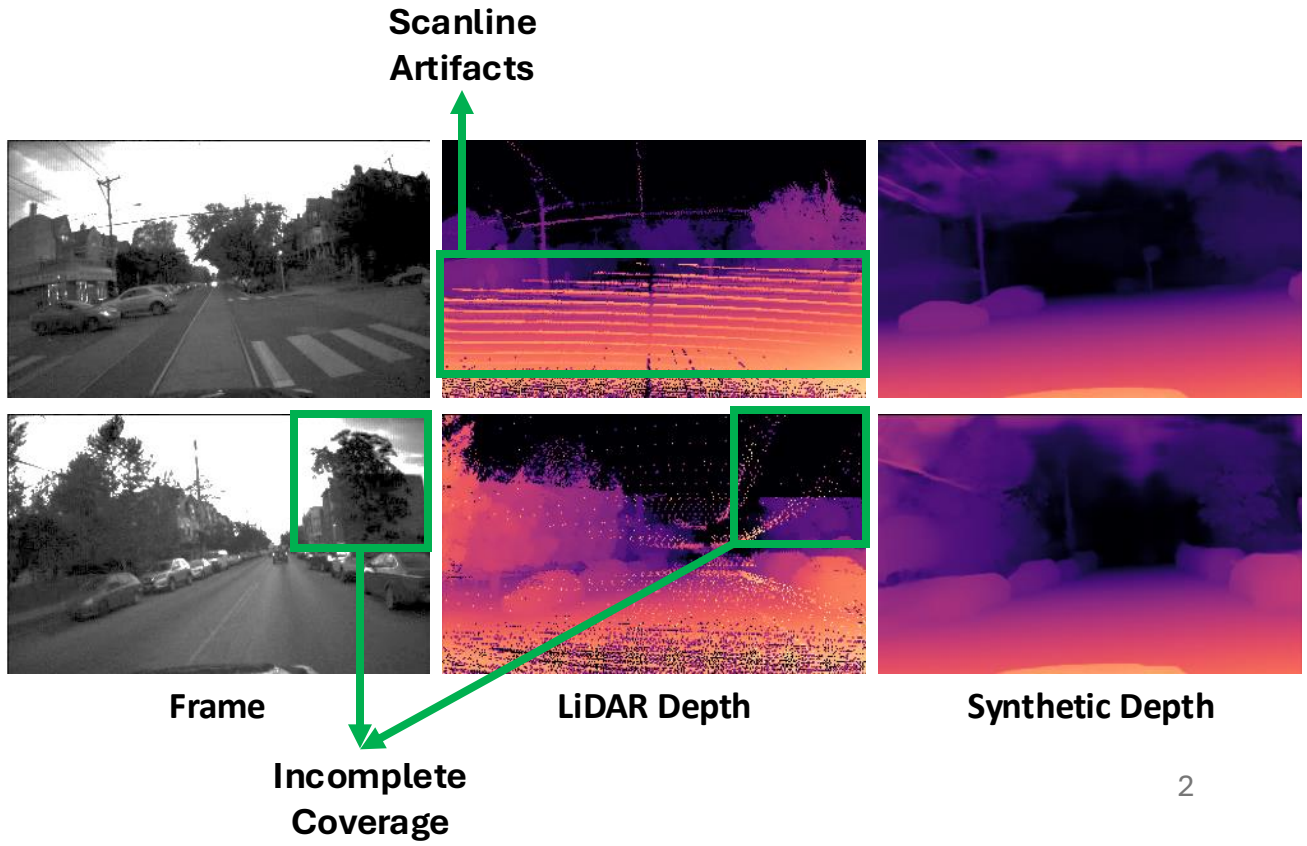
Task: Event-based Monocular Depth Estimation

Constraints: Low quality LiDAR supervision

Idea: Use daytime event + image datasets, generate **synthetic depths** from images using **foundational depth models**

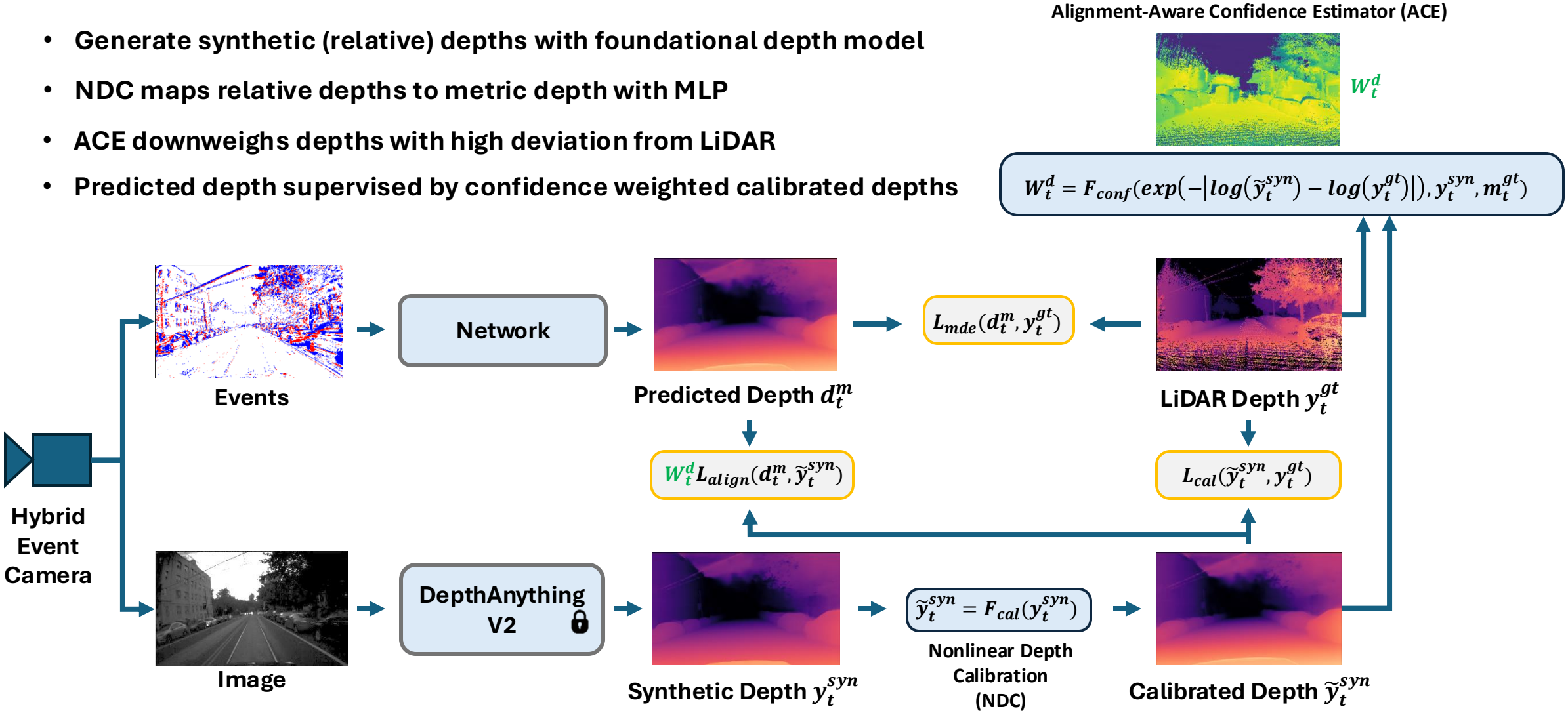


	LiDAR	Synthetic Depth
Dense	✗	✓
Complete Coverage	✗	✓
No Scanline Artifacts	✗	✓
Metric Scale	✓	✗



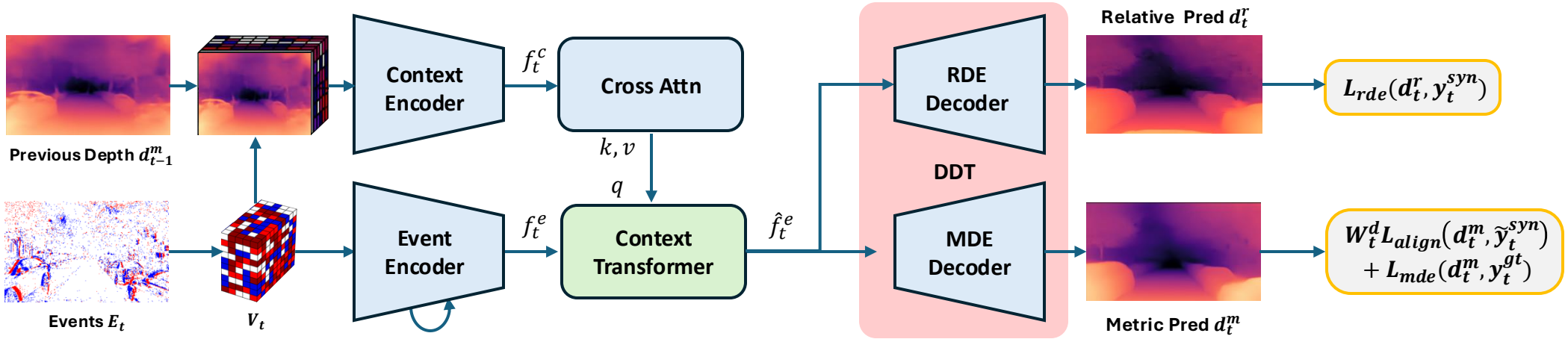
Solution

- Generate synthetic (relative) depths with foundational depth model
- NDC maps relative depths to metric depth with MLP
- ACE downweighs depths with high deviation from LiDAR
- Predicted depth supervised by confidence weighted calibrated depths

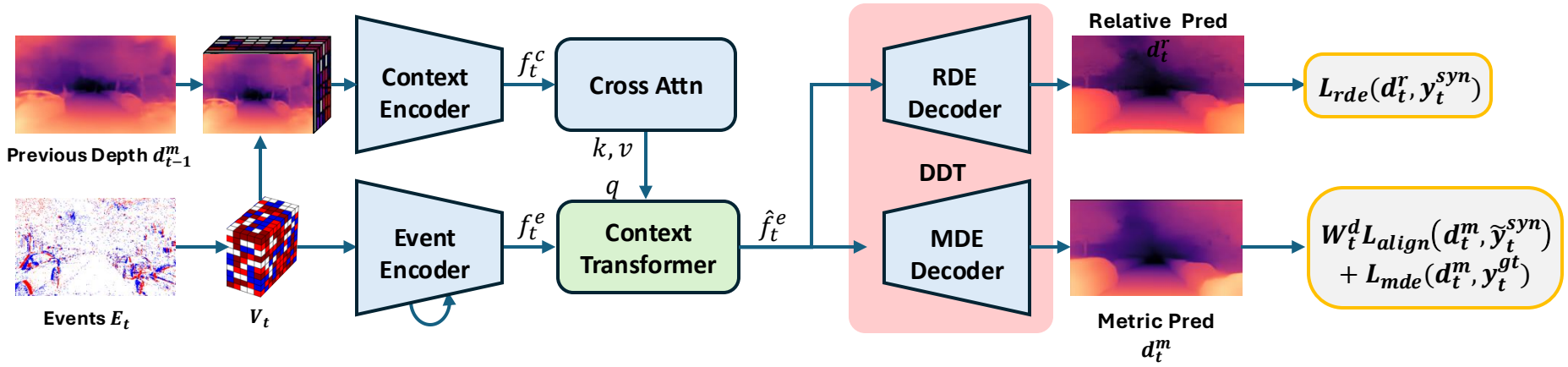


Solution

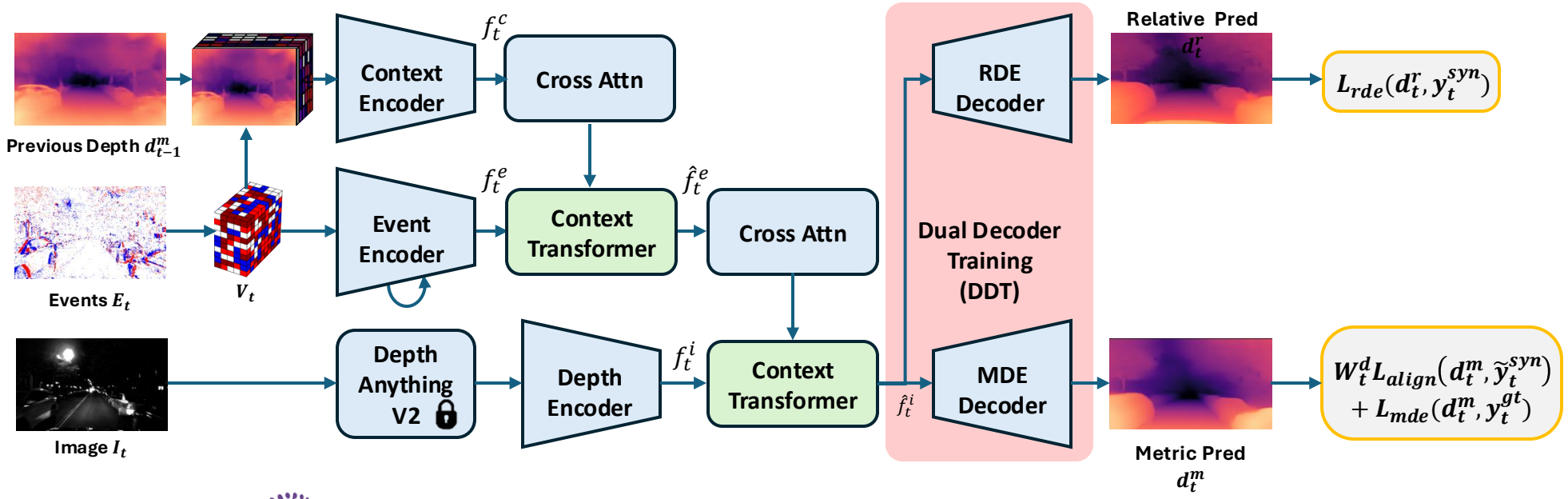
- **Context Transformer:** Incorporates temporal and semantic cues by attending to past depth and event features for consistent depth prediction.
- **Dual Decoder Training:** Employs separate metric and relative depth decoders to specialize learning and jointly regularize the encoder for sharper, more robust depth maps.



Distil-E2D (Event)



Distil-E2D (Event+Image)

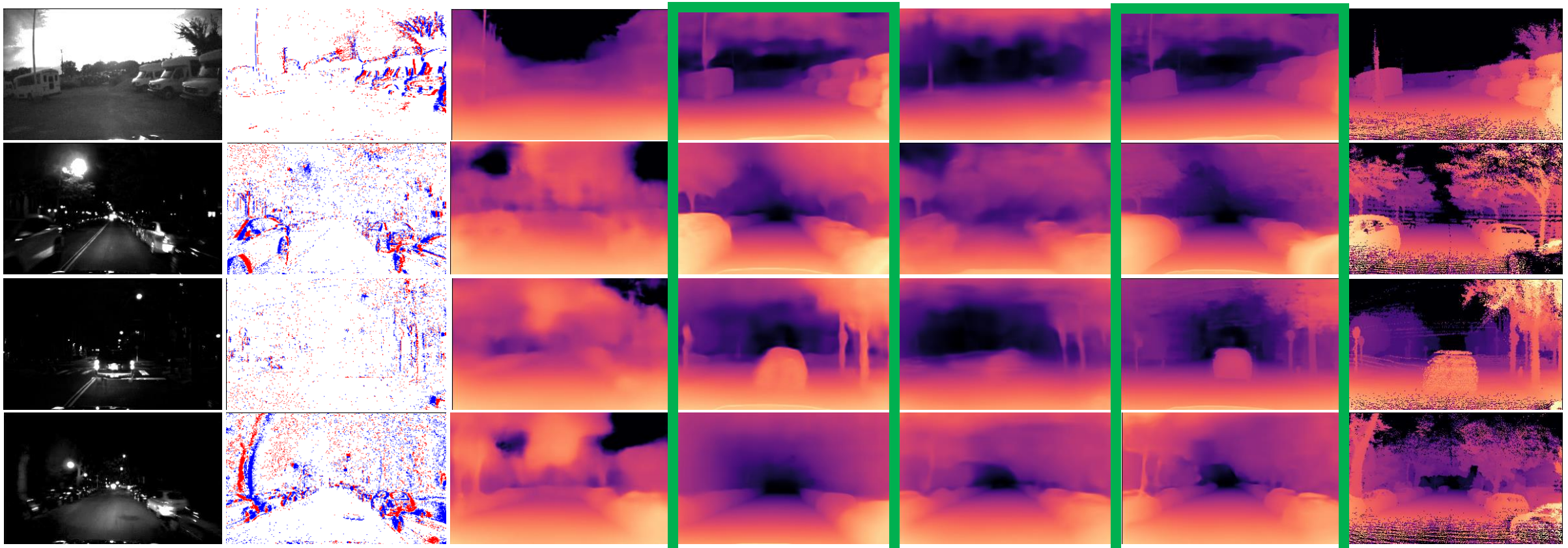


Quantitative Results

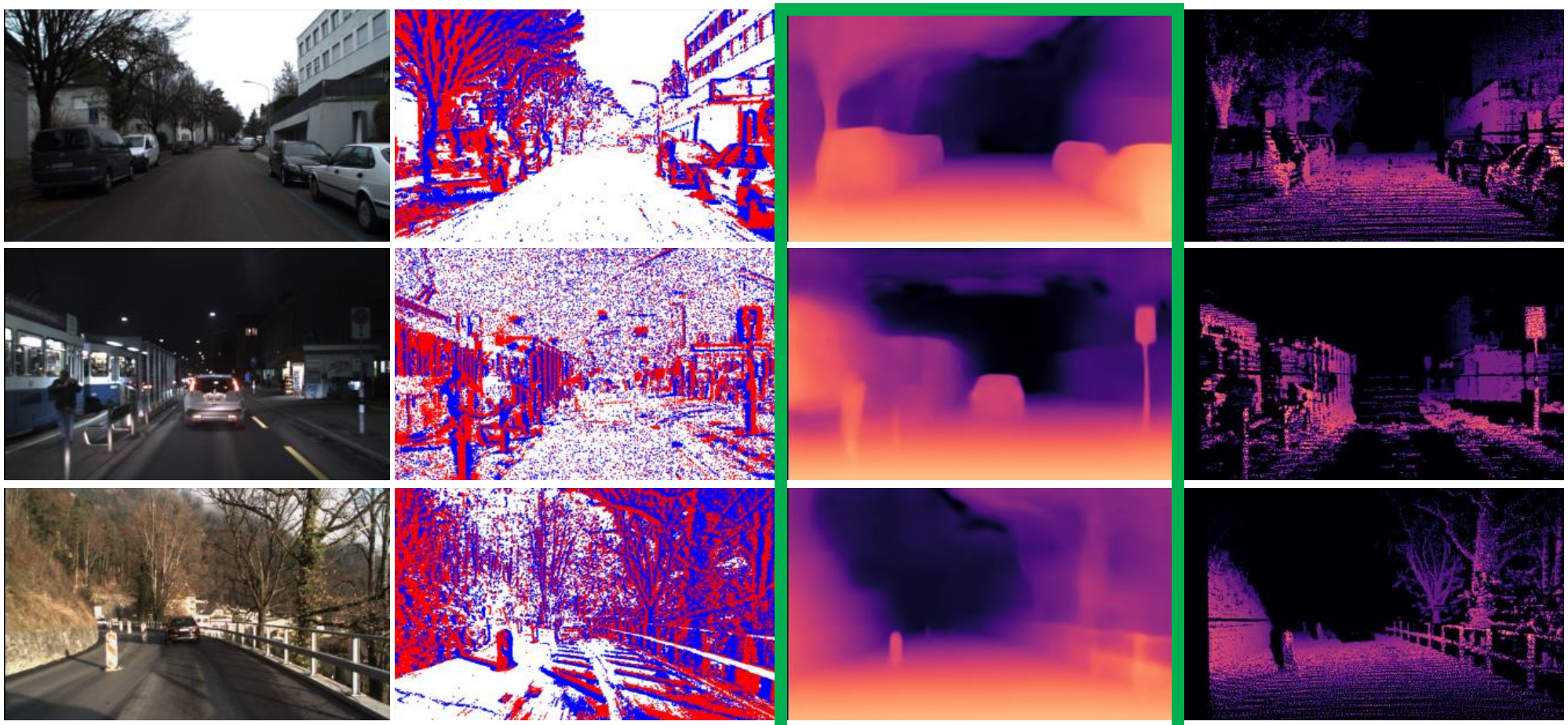
Table 1: Quantitative results on MVSEC Dataset.

	day1			night1			night2			night3			
	10m ↓	20m ↓	30m ↓	10m ↓	20m ↓	30m ↓	10m ↓	20m ↓	30m ↓	10m ↓	20m ↓	30m ↓	runtime(ms)
Event-only (E)													
ULODE[37]	2.72	3.84	4.40	3.13	4.02	4.89	2.19	3.15	3.92	2.86	4.46	5.05	25.0
MDDE+[8]	1.85	2.64	3.13	3.38	3.82	4.46	1.67	2.63	3.58	1.42	2.33	3.18	24.2
EMD[30]	1.40	2.07	2.65	2.18	2.70	3.64	2.06	2.76	3.42	2.09	2.82	3.52	**5.3**
ERE[11]	1.44	2.23	2.75	1.85	2.45	3.24	1.66	2.53	3.15	1.49	2.37	2.97	35.0
EvT+ (E)[12]	1.31	1.92	2.32	1.54	2.31	2.96	1.47	2.22	2.92	1.36	2.13	2.84	35.0
Ours (E)	**1.14**	**1.56**	**1.81**	**1.41**	**2.16**	**2.75**	**1.33**	**2.06**	**2.70**	**1.13**	**2.04**	**2.79**	10.5
Event+Image (E+I)													
RAMNet[9]	1.39	2.17	2.76	2.50	3.19	3.82	1.21	2.31	3.28	1.01	2.34	3.43	**6.2**
EvT+ (E+I)[12]	1.24	1.91	2.36	1.45	2.10	2.88	1.48	2.13	2.90	1.38	2.03	2.77	52.0
Ours (E+I)	**1.04**	**1.44**	**1.75**	**1.35**	**2.02**	**2.79**	**1.20**	**2.04**	**2.62**	**0.99**	**1.96**	**2.71**	47.5

Qualitative Results (MVSEC)



Qualitative Results (DSEC)



Frame | Events | Distil-E2D (E) (Ours) | GT Depth