ICCV OCT 19-23, 2025 HONOLULU HAWAII

Enhanced Event-based Dense Stereo via Cross-Sensor Knowledge Distillation

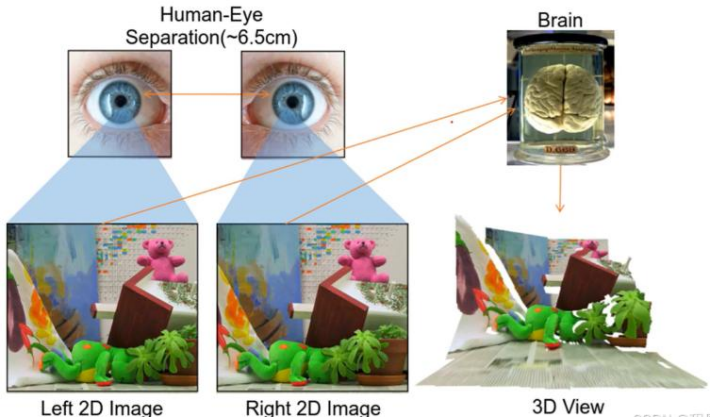
Haihao Zhang, Yunjian Zhang✉, Jianing Li✉, Lin Zhu, Meng Lv, Yao Zhu, Yanwei Liu, Xiangyang Ji

Institute of Information Engeering, Chinese Academy of Sciences, Beijing Institute of Technology, Tsinghua University

Exhibit Hall I #500

E-mail: haihaozhang@bit.edu.cn, sdtczyj@gmail.com, lijianing@pku.edu.cn

Stereo Matching

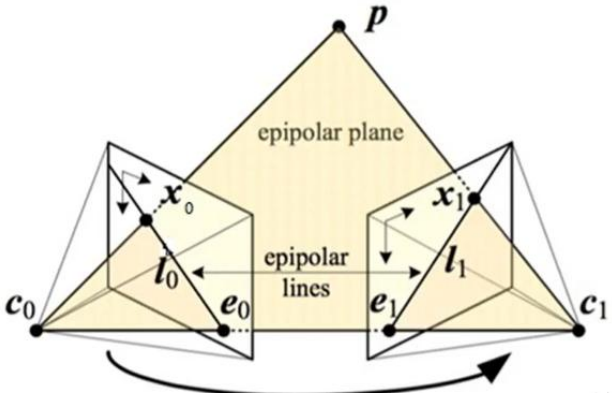


Human binocular perception

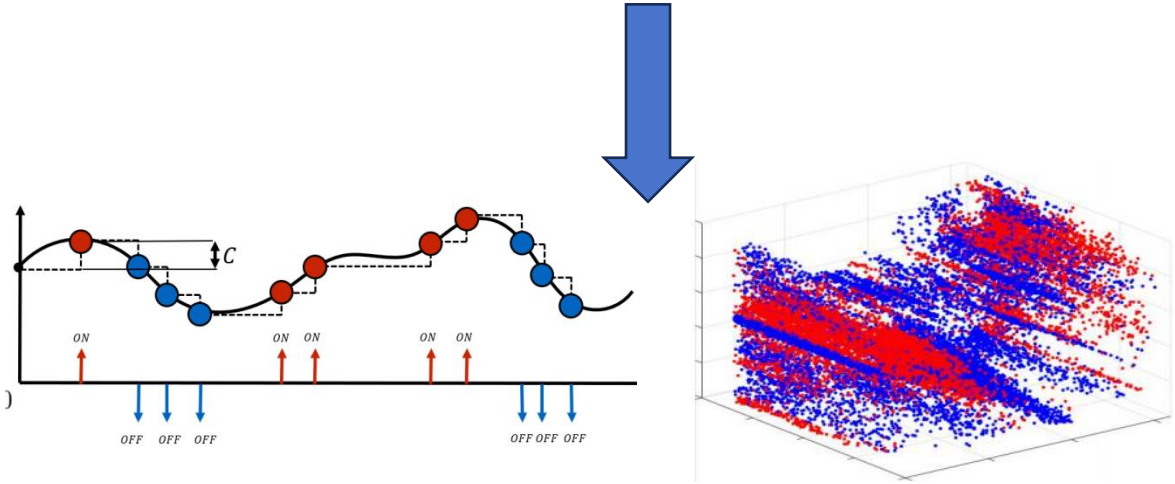
RGB's Challenges

Motion Blur

Limited dynamic range



Stereo matching



Low latency
High dynamic range

Event-based Camera

Event-based Methods

Two Categories

Event Only

Advantages:

Low latency inference (due to the characteristics of events)

Limitations

The sparseness of input events increases the difficulty of obtaining a dense disparity map.

Event & RGB

Advantages:

Achieves the best performance by leveraging complementary information between two modalities.

Limitations

Simultaneously inputting two modalities during inference sacrifices the low latency of events and increases the inference burden.

Motivation

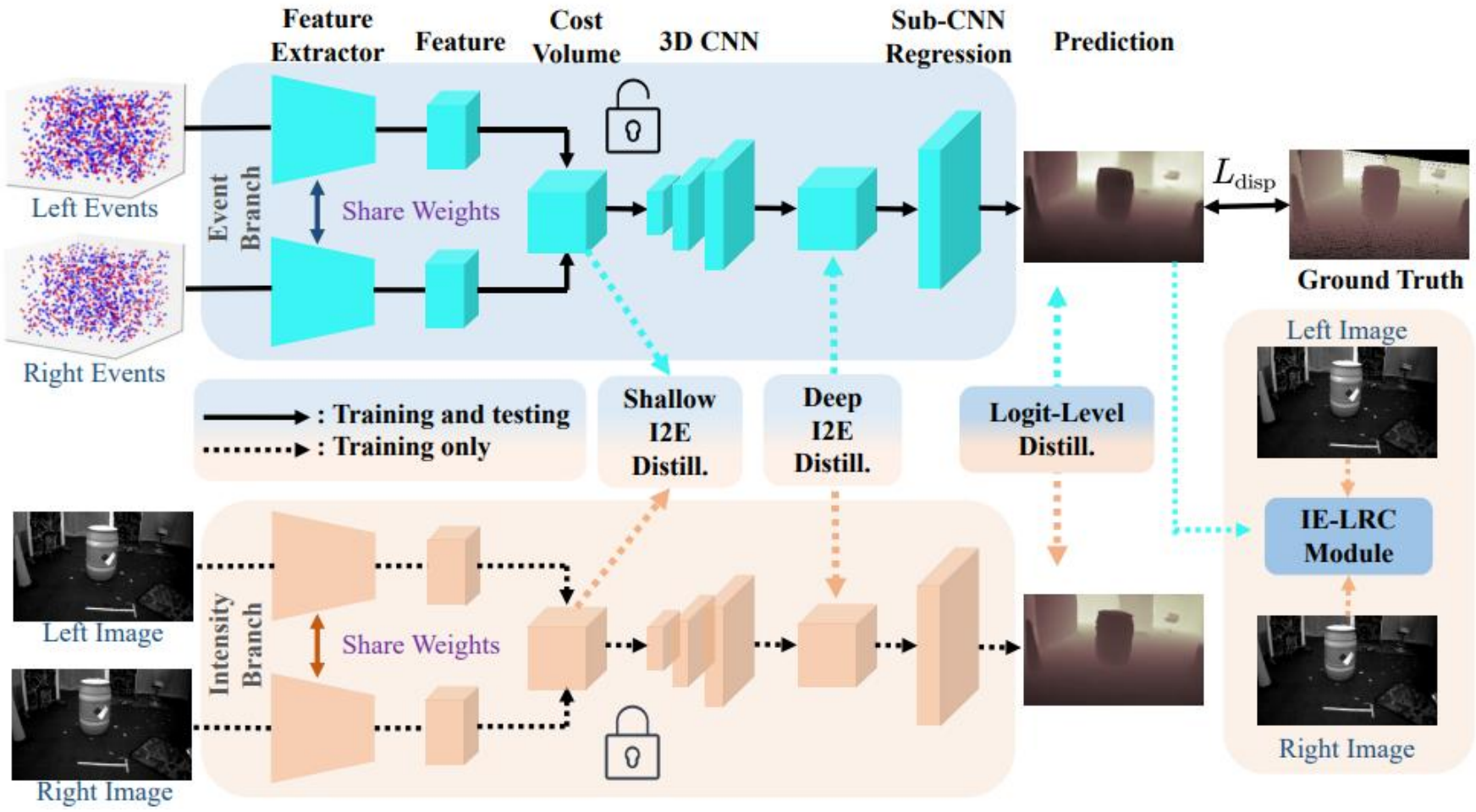
Cross-sensor Distillation

- During training, two modalities are input to fully capture information from RGB and di still it to events.
- While during inference, only events are input. This preserves the low latency of events.

Challenges

- Gap between different modalities. → Multi-level distillation strategy
- Possible loss imbalance. → Distill the cost volume rather than the feature

Approach

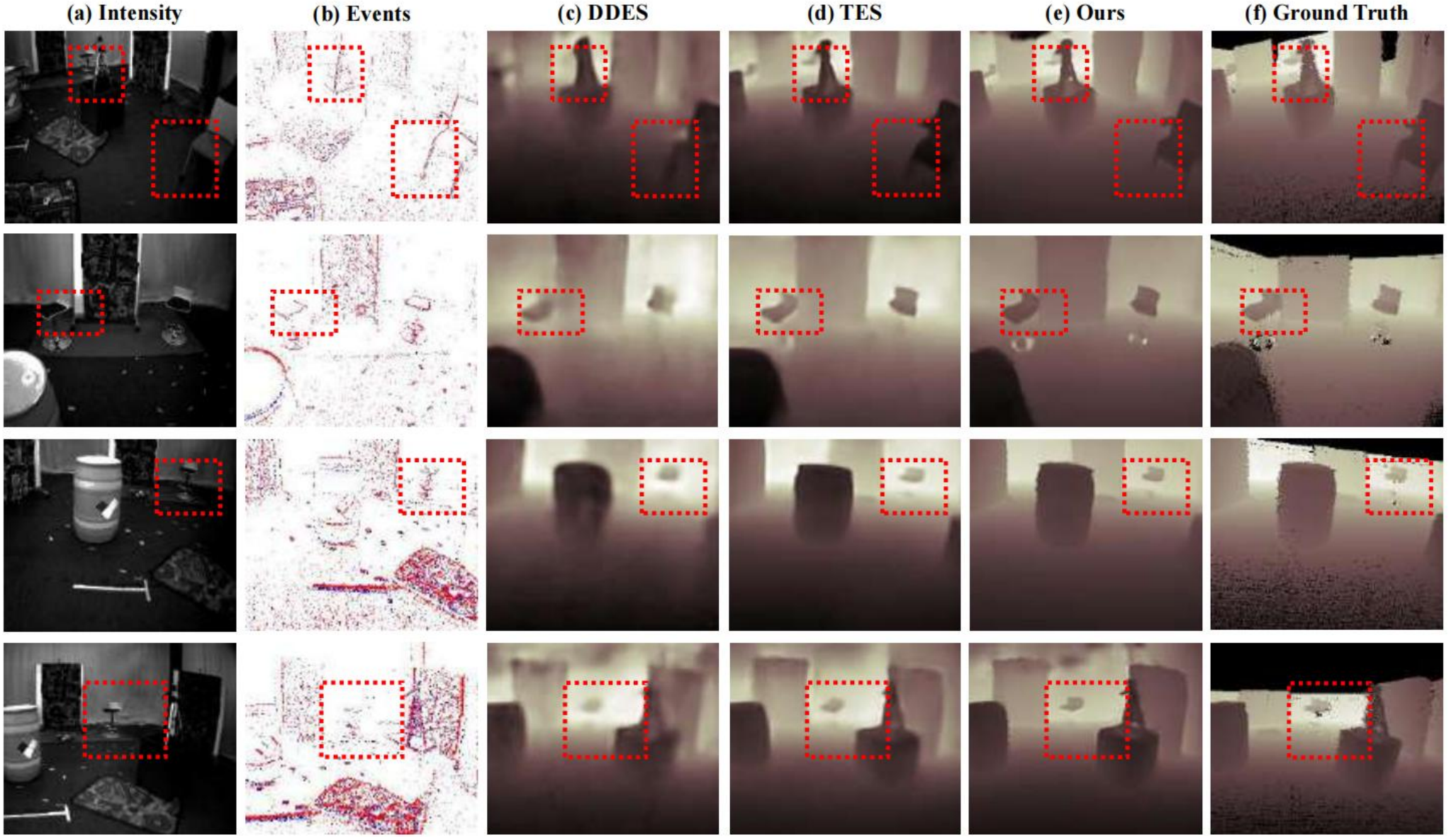


The multi-level distillation strategy considers both the shallow features related to the modality and the deep features related to the task.

The distillation is based on the cost volume, which eliminates the loss imbalance problem.

Event-RGB left-right consistency module to enhance structural consistency.

Results - MVSEC



Results - Ablations

F-distill.	L-distill.	LRC	Mean disparity error [pix] ↓		One-pixel accuracy [%] ↑		Mean depth error [cm] ↓	
			Split 1	Split 3	Split 1	Split 3	Split1	Split3
✗	✗	✗	0.54	0.67	91.0	89.4	15.0	18.5
✓	✗	✗	0.38	0.46	94.4	92.0	12.1	13.9
✗	✓	✗	0.49	0.53	91.9	91.6	13.8	17.5
✗	✗	✓	0.50	0.64	91.2	90.3	14.6	18.0
✗	✓	✓	0.49	0.52	92.1	91.6	13.7	17.4
✓	✗	✓	0.37	0.41	94.5	92.7	11.6	13.7
✓	✓	✗	0.34	0.39	94.8	93.2	11.3	13.4
✓	✓	✓	0.32	0.37	94.8	93.4	11.2	13.2

Methods	Mean disparity error [pix] ↓		One-pixel accuracy [%] ↑	
	Split1	Split3	Split1	Split3
MSE	0.45	0.50	92.2	91.8
Si2eD	0.39	0.46	93.0	92.5
MSE+Di2eD	0.37	0.41	94.0	93.0
Si2eD+Di2eD	0.32	0.37	94.8	93.4

Thanks for your attention