Dense Depth from Event Focal Stack

Kenta Horikawa[1] Mariko Isogawa[1] Hideo Saito[1] Shohei Mori[2,1]
[1]Keio University [2]University of Stuttgart
University of Stuttgart Germany

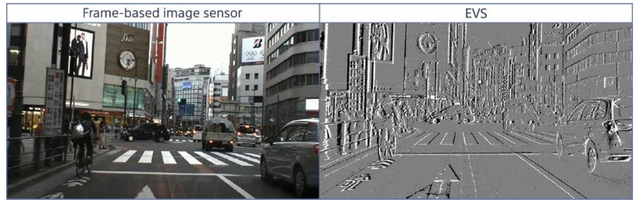
WACV 2025 TUCSON, ARIZONA • FEB 28 - MAR 4
Paper ID: 2203

Overview

- Depth estimation which is robust to extreme light conditions (i.e., low-light) and be able to utilize in static scenes
- Propose dense depth estimation only from event focal stack
  - 3D representation of the event sequence triggered by focal sweep
- Own synthetic & real-captured dataset
- Our framework shows better accuracy and better robustness to the low-light conditions than image-based method

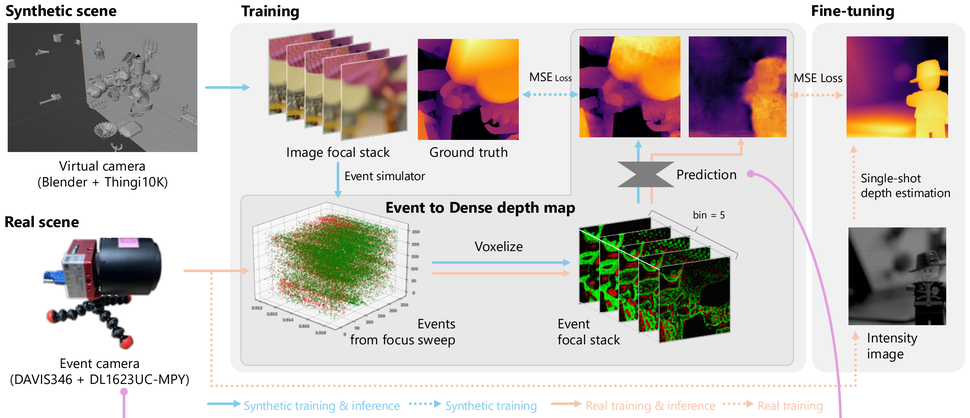
Background

- Limitation of image-based method
  - fails in extreme light conditions
- Event-based camera
  - Event: records brightness changes asynchronously
  - High temporal resolution, HDR



Proposed Method

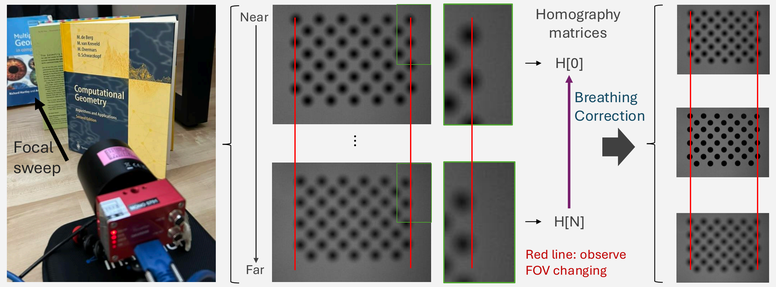
- Events from focal sweep → Dense depth map
- Train on synthetic data, then finetune with real-captured data



- The event camera equipped with a computer-controlled lens

U-Net like architecture derived from event to video method[1]

- Breathing correction with homography matrices

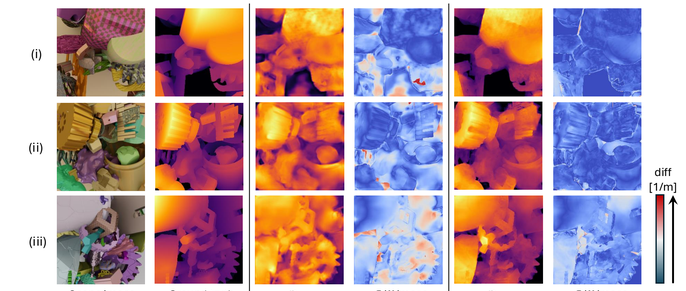


Experimental Results

- Result of synthetic data

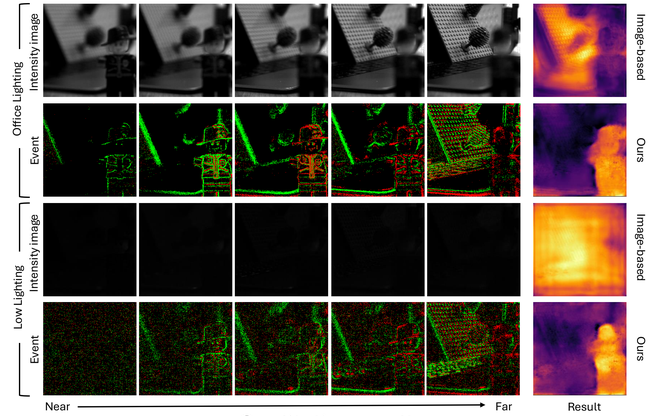
Quantitative result

Method	MAE[1/m](↓)	RMSE[1/m](↓)
Focus on Defocus[2]	0.1606	0.2027
Ours	**0.0762**	**0.1022**



Qualitative result

- Result of real-captured data



Qualitative result

[1] Henri Rebecq *et al.*, "High Speed and High Dynamic Range Video with an Event Camera". TPAMI2019
[2] Maxim Maximov *et al.*, "Focus on Defocus: Bridging the Synthetic to Real Domain Gap for Depth Estimation". CVPR2020