Yuda DONG, Zetao CHEN, Xin HE, Lijun LI, Zichao SHU, Yinong CAO, Junchi FENG, Shijie LIU, Chunlai LI, Jianyu WANG, 2024. SEVAR: a stereo event camera dataset for virtual and augmented reality. *Frontiers of Information Technology & Electronic Engineering*, 25(5):755-762. https://doi.org/10.1631/FITEE.2400011

SEVAR: a stereo event camera dataset for virtual and augmented reality

Key words: Event camera; Virtual and augmented reality (VR/AR); Time synchronization and calibration; Simultaneous localization and mapping (SLAM); Dataset

Corresponding authors: Zetao CHEN, Xin HE
E-mail: zetao-chen@ylab.ac.cn; xinhe@ucas.ac.cn
ORCID: Zetao CHEN, https://orcid.org/0000-0002-5596-5008; Xin HE, https://orcid.org/0000-0001-8229-8143

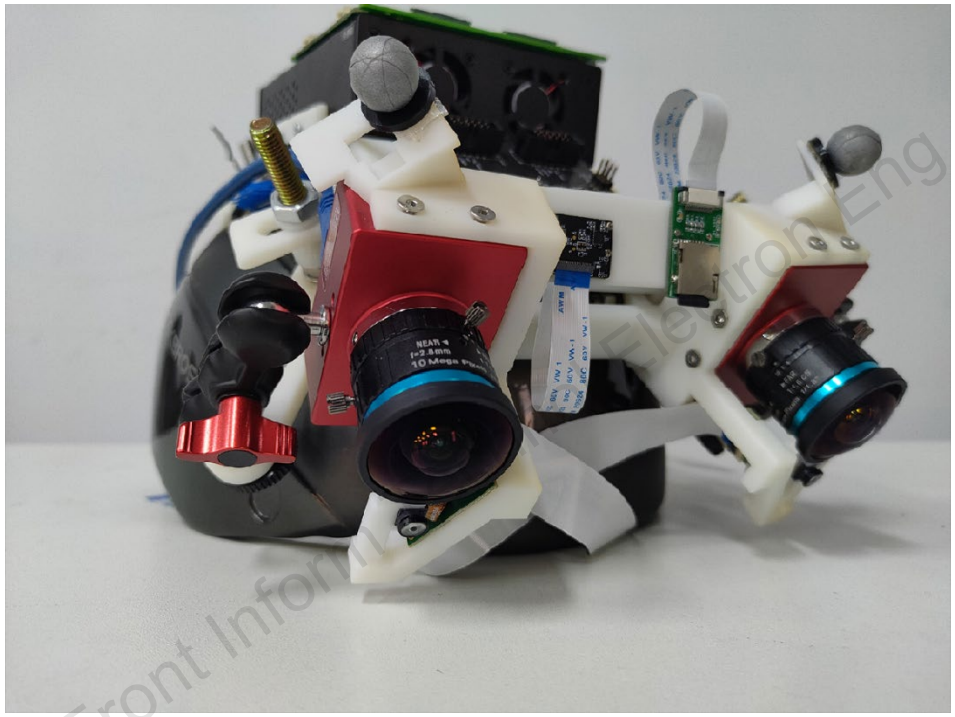
Motivation

Event cameras, characterized by their low latency, large dynamic range, and extremely high temporal resolution, have recently received increasing attention. These features make them particularly well-suited for virtual/augmented reality (VR/AR) applications. To facilitate the development of three-dimensional (3D) perception and navigation algorithms in VR/AR applications using event cameras, we introduce the Stereo Event camera dataset for Virtual and Augmented Reality (SEVAR), which comprises a wide variety of head-mounted indoor sequences, including scenarios with rapid motion and a large dynamic range.

Main idea

We present the first comprehensive set of VR/AR datasets captured with an event-based stereo camera, a regular stereo camera at 30 Hz, and an inertial measurement unit (IMU) at 1000 Hz. The camera placement, field of view, and resolution match those of the head mounted device, such as Meta Quest Pro. All sensors are time-synchronized in the hardware. Ground truth poses captured by a motion capture system are also available for trajectory evaluation. The sequences include several common scenarios, and cover the specific challenges targeted by event cameras.

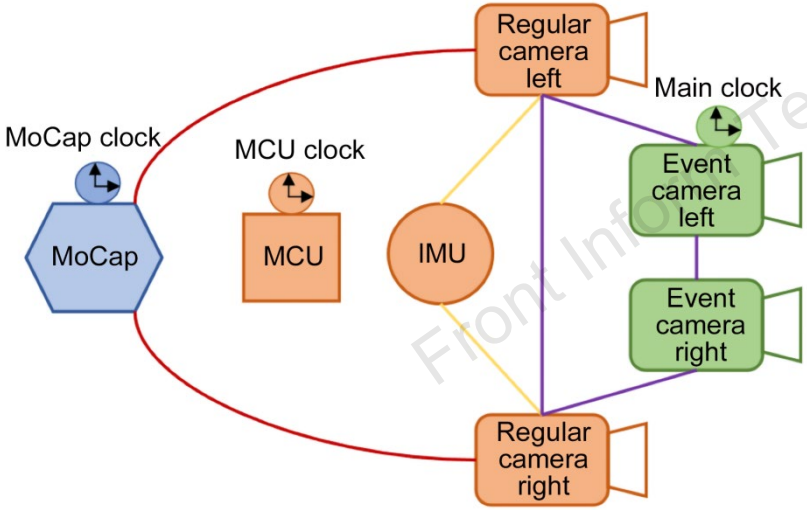
Framework



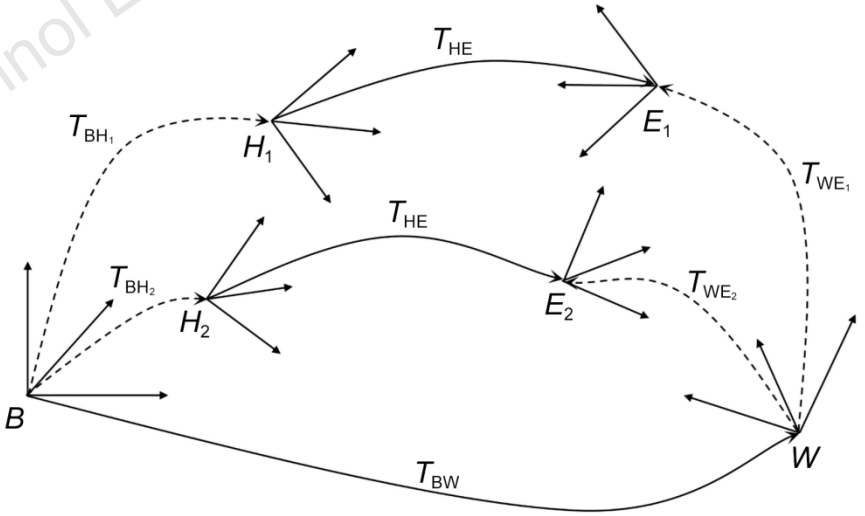
Overview of the acquisition platform. In this configuration, all sensors, including the IMU, regular cameras, and event cameras, are rigidly mounted on a 3D printed holder. The configuration and technical specifications of the regular cameras and IMU closely resemble those found in commercial VR devices, with the event cameras being positioned directly above this arrangement.

Method

Building the acquisition platform consists of two tasks. The first task is time synchronization. The time synchronization of all sensors is achieved through the use of hardware triggers provided by a micro-controller unit (MCU). The second task is calibration.

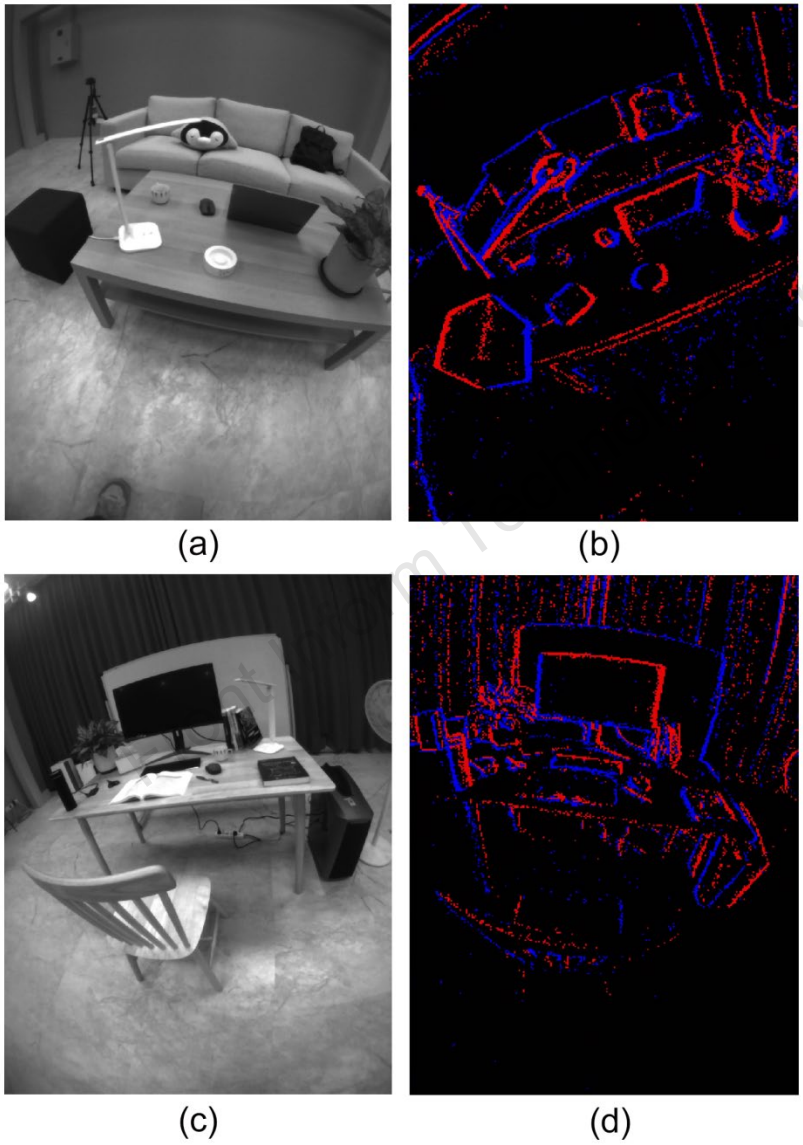


Overview of time synchronization



Overview of calibration

Major sequences

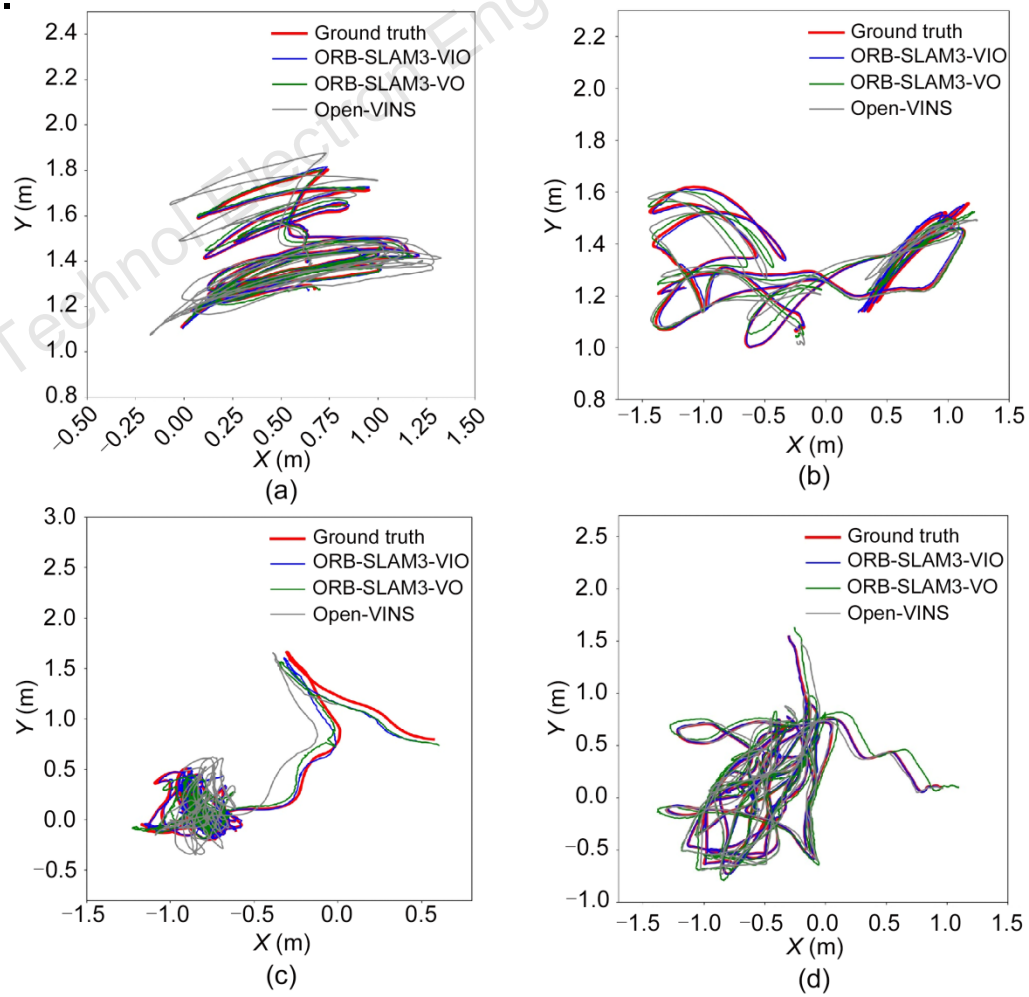


Examples of sofa-fast and desk-normal sequences:
(a) sofa-frame data;
(b) sofa-event data;
(c) desk-frame data;
(d) desk-event data

Major results

To verify the quality of our data and calibration, we tested the sequences using state-of-the-art open-source algorithms, including ORB-SLAM3 and Open-VINS.

Visual trajectory maps for different sequences:
(a) desk-normal;
(b) sofa-fast;
(c) VR-normal;
(d) AR-normal



Conclusions

We present a precisely synchronized event-based dataset, designed especially for multi-sensor fusion in SLAM applications, with a particular emphasis on VR and AR scenarios. Alongside setting up commonly used stereo regular cameras and an IMU, we have integrated stereo event cameras. We specialize in recording sequences to imitate real-life scenarios, while adding challenging sequences such as low light and fast motion. It is our aspiration that this dataset will serve as a valuable resource for the advancement of research in the domain of event-based multi-sensor fusion algorithms. The dataset can be found at https://github.com/sevar-dataset/sevar.