# Optical Flow-Guided 6DoF Object Pose Tracking with an Event Camera

Anonymous Authors

## 1 EXPERIMENTAL PREPARATION

Due to the lack of an appropriate event dataset for 3D object motion, we prepare simulated and real events to evaluate the performance of the proposed method. In this section, we present additional details regarding the experimental preparations.

### 1.1 Simulated events

We select two representative categories of objects: those with straight edges, and those with curved edges from [1], effectively encompassing a wide range of common objects. These objects undergo substantial rotations and translations along various motion trajectories, while being captured using a stationary monocular camera at a frame rate of 30 Hz and a resolution of 640×480 pixels. The resulting sequence of RGB videos is rendered using Blender. This process employs Blender to accomplish photorealistic rendering and generate a series of RGB videos. Then, these videos are converted into event streams using V2E [2], while keeping the camera and video parameters unchanged.

**Object with straight edges.** The rendered RGB images and corresponding simulated events of objects with straight edges are depicted in Figure 1. Above are the rendered RGB images, while below are the corresponding event accumulated images. These objects move along predetermined trajectories. They are classified into fast, normal, and slow speeds based on their respective velocities. In the experiment, we conduct accuracy validations separately for the aforementioned scenarios.

**Object with curved edges.** To simulate various complex scenarios encountered during object tracking, we introduce several challenging factors, including cluttered backgrounds and extensive occlusions. The rendered RGB images and corresponding simulated events are shown in Figure 2. Left side of Figure 2 illustrates rendered images of moving objects against a complex background. In such scenarios, traditional monocular tracking algorithms may be susceptible to the influence of color similarity and background complexity, resulting in a decline in tracking performance. However, for event cameras, stationary objects scarcely generate events. Despite the minimal interference from background events generated by v2e, our method continues to effectively maintain stable tracking of objects. The right side of Figure 2 illustrates situations where objects are occluded, presenting a challenging scenario. Our method, leveraging optical flow guidance to establish associations

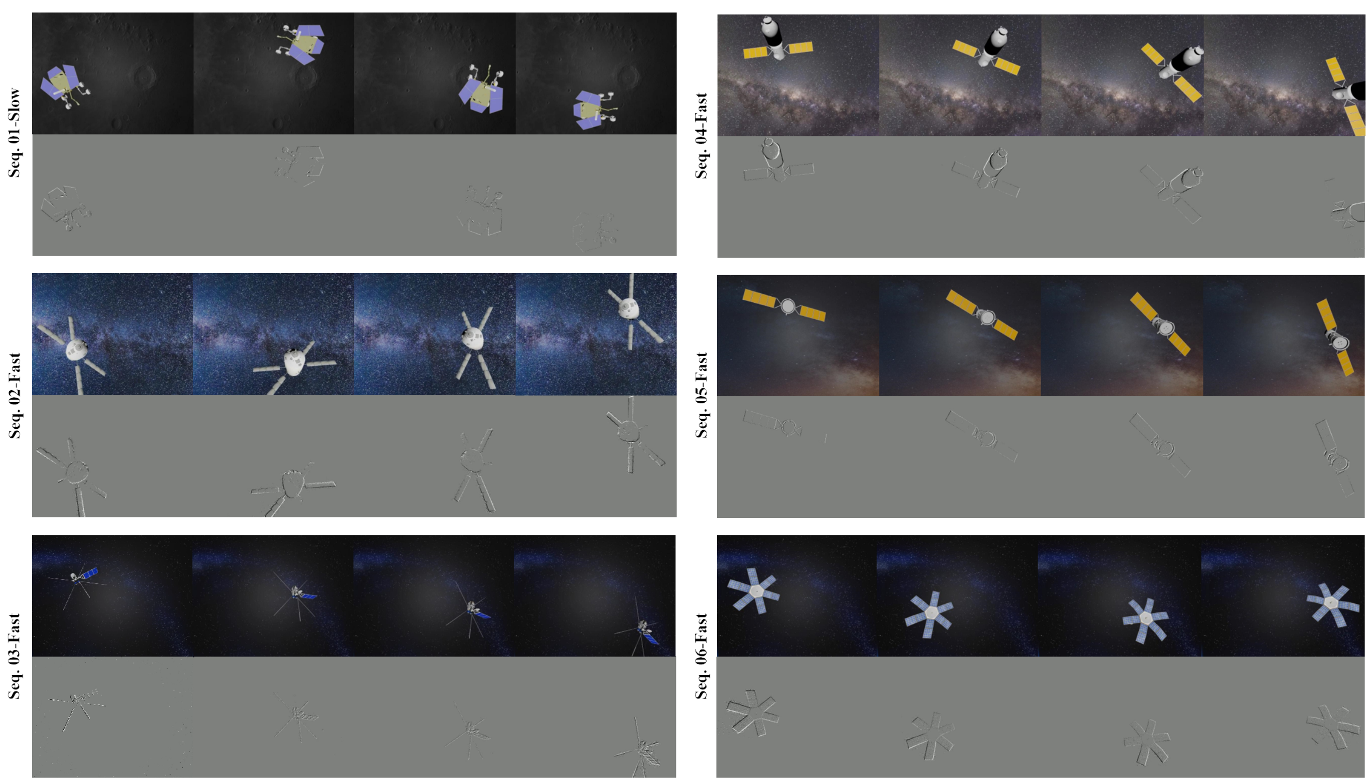


**Figure 1: The rendered RGB images and corresponding simulated events of objects with straight edges.**

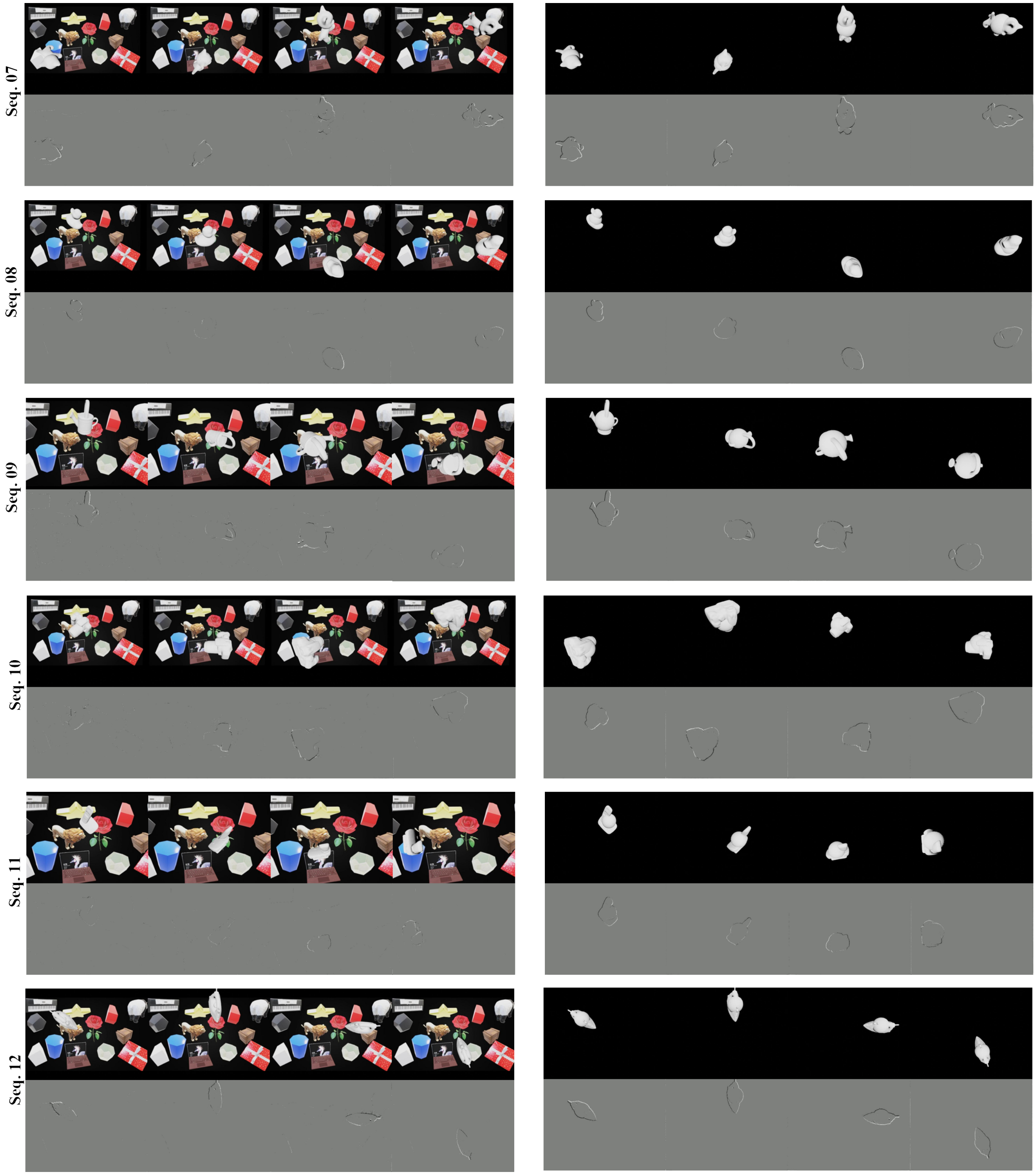


**Figure 2: The rendered RGB images and corresponding simulated events of objects with curved edges.**

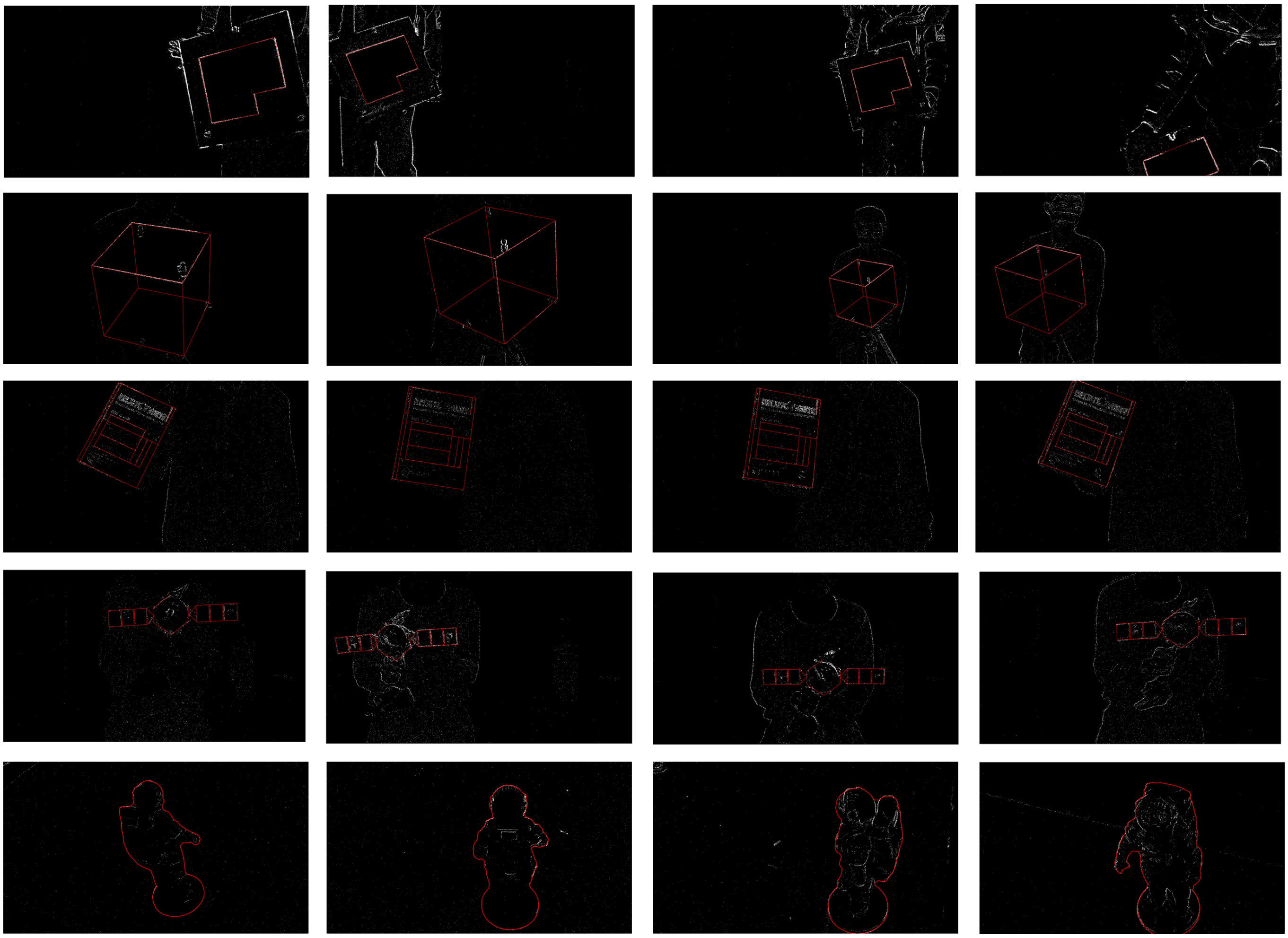

**Figure 3: Pose tracking results of real event experiments. The red represents the edges of the 3D model, which are reprojected onto accumulated event images using estimated poses obtained from our methods.**

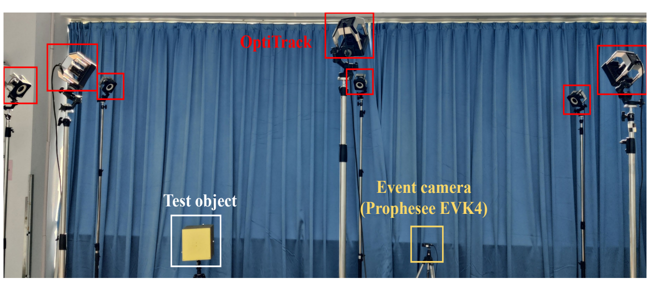


**Figure 4: Experimental Scene Setup. Objects in motion are captured by a stationary event camera, with pose ground truth provided by OptiTrack.**

between events and features, effectively tackles the aforementioned challenges, achieving high-precision pose tracking.

## 1.2 Real events.

To validate the feasibility of the proposed method, we conduct multiple tests using real events, including objects with straight and curved edges. The models of these objects are known a priori. Before conducting the tests, we attach markers to the surface of objects, which are used to obtain the ground truth pose of objects using the OptiTrack system. The experimental scenario is as depicted in Figure 4. Firstly, we manually control the motion of objects and record their movement using a pre-calibrated and fixed event camera (Prophesee EVK4, resolution 1280 × 720 pixels). Subsequently, the tracking methods are validated using event streams and compared against other advanced algorithms. For the astronaut model, we control the camera movement while keeping the model stationary, thereby enabling a more comprehensive test of the method.

The pose tracking results are visually displayed in Figure 3. The projected edges of the astronaut model remain tightly aligned with events throughout the entire tracking process, providing an intuitive demonstration of the accuracy of our tracking methods.

## REFERENCES

[1] Stefan Hinterstoisser, Vincent Lepetit, Slobodan Ilic, Stefan Holzer, Gary Bradski, Kurt Konolige, and Nassir Navab. 2012. Model based training, detection and pose estimation of texture-less 3d objects in heavily cluttered scenes. In *Proc. Asian Conf. Comp. Vis.* (Daejeon, Korea). 548–562.

[2] Yuhuang Hu, Shih-Chii Liu, and Tobi Delbrück. 2021. v2e: From Video Frames to Realistic DVS Events. *IEEE Conf. Comput. Vis. Pattern Recognit. Work.* (2021), 1312–1321.