Robust Hand-Object Pose estimation from Multi-view 2D Keypoints

Hand-object pose estimation is a challenging task due to multiple factors like occlusion, and ambiguity in pose recovery. To overcome this issue, multi-view camera systems are used.

Using 2D keypoint detectors for hands and objects like Yolov8-pose and mmpose we can uplift the 2D detections to 3D. However, the detections usually are usually noisy, and some keypoints may be missing.

We want to utilize deep learning methods for smoothing, inpainting, and uplifting these detections to 3D in order to estimate the pose of the corresponding hands and objects.

The task is formulated as follows:

Given a sequence of noisy 2D key points for human hands and an object captured from calibrated camera views. Using a deep learning model, estimate a smooth trajectory of the hand and object poses.

Prerequisites

- Python
- Knowledge about Deep Learning
- Knowledge about Pytorch
- Previous Knowledge about 3D data processing is a plus.

Contact

marsil.zakour@tum.de

Advisors

Marsil Zakour