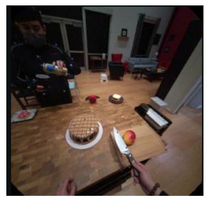
Power-Efficient SLAM for AR Glasses

RUTGERS
WINLAB | Wireless Information Network Laboratory

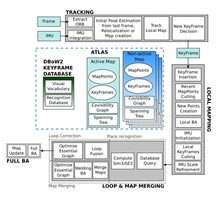
Intern: Easton Wei
Advisors: Prof. Yao Liu, Jason Ding

Project Overview

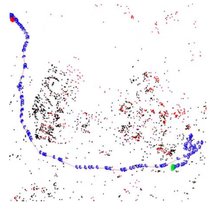
- Augmented Reality (**AR**) requires efficient, continuous processing of visual data
- In the realm of AR, Simultaneous Localization and Mapping (**SLAM**) is required to track the user's position and anchor virtual objects in the physical world



Input frame



SLAM pipeline [1]

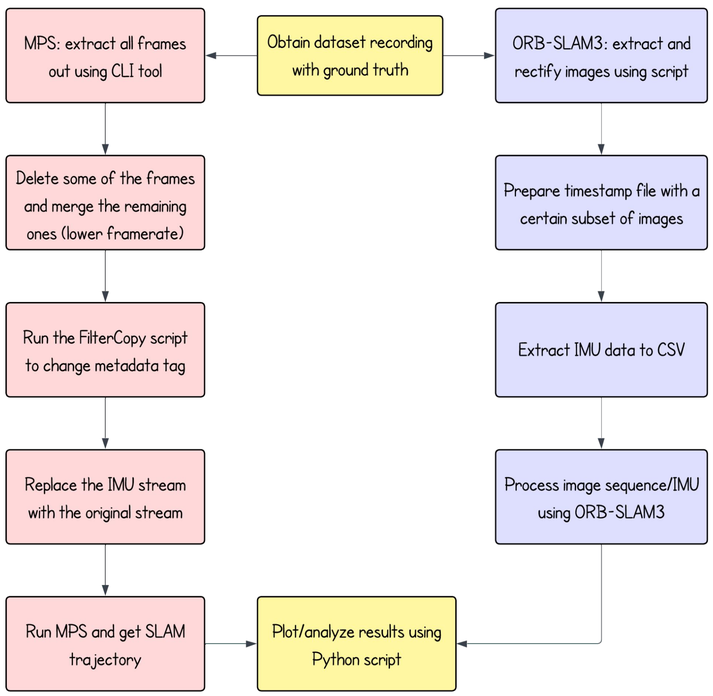


Output trajectory and pointcloud

- One challenge of SLAM for AR is that on-device data collection requires significant amounts of power
- Two potential ways to decrease power usage:
  - Decreasing the framerate (focus of the project)
  - Decreasing the resolution

Experimental Design

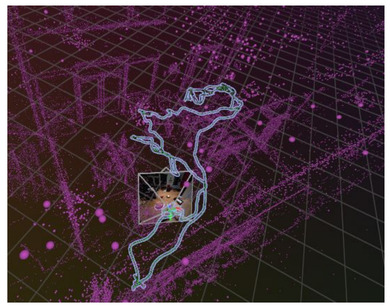
- This experiment aims to profile how different framerates affect the accuracy of SLAM outputs
- Recordings and ground truth from the Aria Digital Twin dataset [2] were used during the project



[1]: https://github.com/UZ-SLAMLab/ORB_SLAM3
[2]: https://www.projectaria.com/datasets/adt/
[3]: https://github.com/facebookresearch/egocentric_splats/
[4]: https://facebookresearch.github.io/projectaria_tools/docs/tech_insights/camera_intrinsic_models

Analysis of MPS

- Aria Machine Perception Services (MPS) is a proprietary SLAM developed by Meta for processing egocentric data captured by Aria glasses

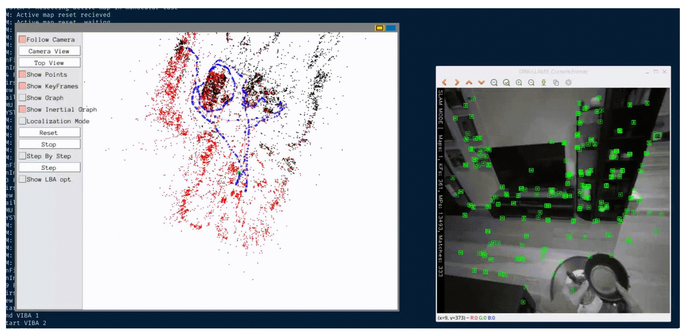


Visualization of MPS output

- MPS uses VRS, which is a file format specifically designed for Aria data streams
- Frames needed to be extracted one-at-a-time from original VRS recordings
- A metadata tag was also edited
- MPS had an overall RMS ATE of **3-5 mm**

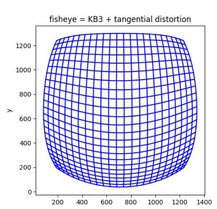
Analysis of ORB-SLAM3

- ORB-SLAM3 [1] is an open-source SLAM library that supports monocular and stereo SLAM, optionally with inertial measurement unit (IMU) data

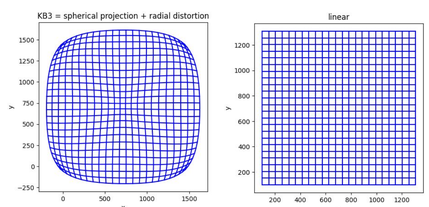


Screenshot of ORB-SLAM3 running

- The camera calibration for Aria glasses is not natively supported by ORB-SLAM3
- Code from "Photoreal Scene Reconstruction from an Egocentric Device" [3] was used to rectify frames
- ORB-SLAM3 had an overall RMS ATE of **2-6 cm**



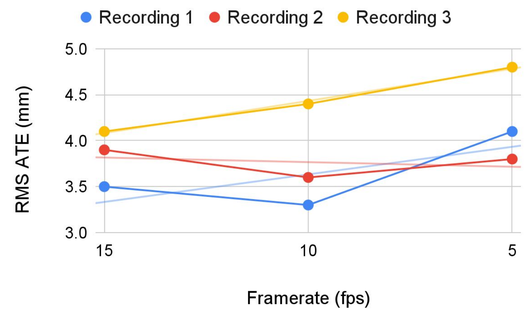
Aria camera calibration (without thin-prism distortion) [4]



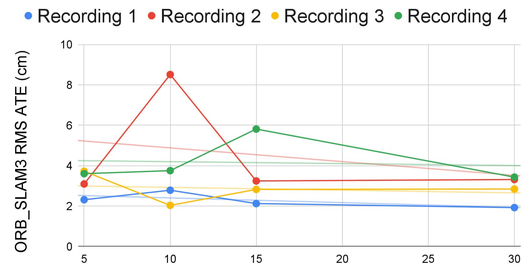
ORB-SLAM3 supported calibrations [4]

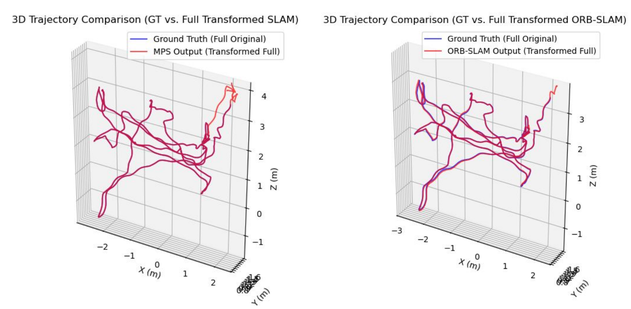
Results

- MPS was one order of magnitude better than ORB-SLAM3 in terms of RMS ATE
- Even at the lowest framerates tested (5fps for MPS and 3fps for ORB-SLAM3), both SLAMs remain accurate enough for AR applications



Performance of MPS (top) and ORB-SLAM3 (bottom) on multiple sequences





Plots of MPS (left) and ORB-SLAM3 (right) trajectories vs. ground truth

Future Work

- Perform quantitative power profiling on Aria glasses and similar devices
- Create a real-time ORB-SLAM3 pipeline to evaluate live performance

Acknowledgements

- I would like to thank Prof. Yao Liu and Jason Ding for their invaluable advice and guidance throughout all stages of this project

WINLAB