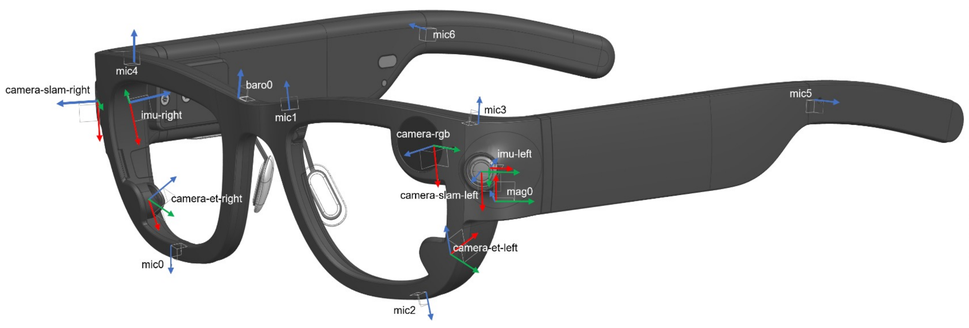
Sparse Image and Dual-IMU Localization for AR Glasses

Students: Dishani Lahiri, Rutika Moharir | S14 Advisor: Prof. Kris Kitani

Problem Statement: Develop a low-power state estimation algorithm for Aria glasses taking a history of IMU measurements and sparse number of camera images as input.

Problem Setting:



Pose Estimates:

$$Position\ p: p_x, p_y, p_z$$

$$Orientation\ q: q_x, q_y, q_z$$

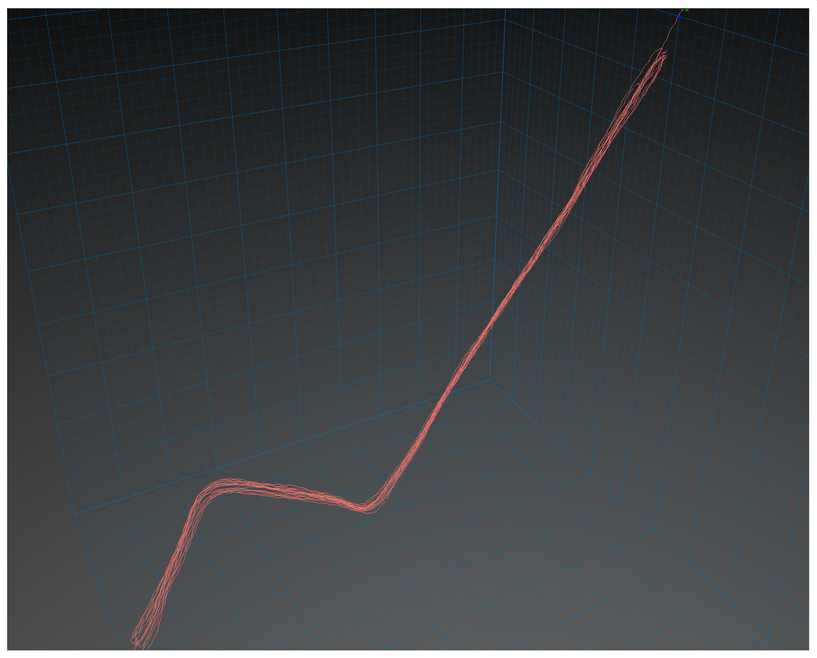
IMU Readings:

$$Accelerometer\ a: a_x, a_y, a_z$$

$$Gyroscope\ \omega: g_x, g_y, g_z$$

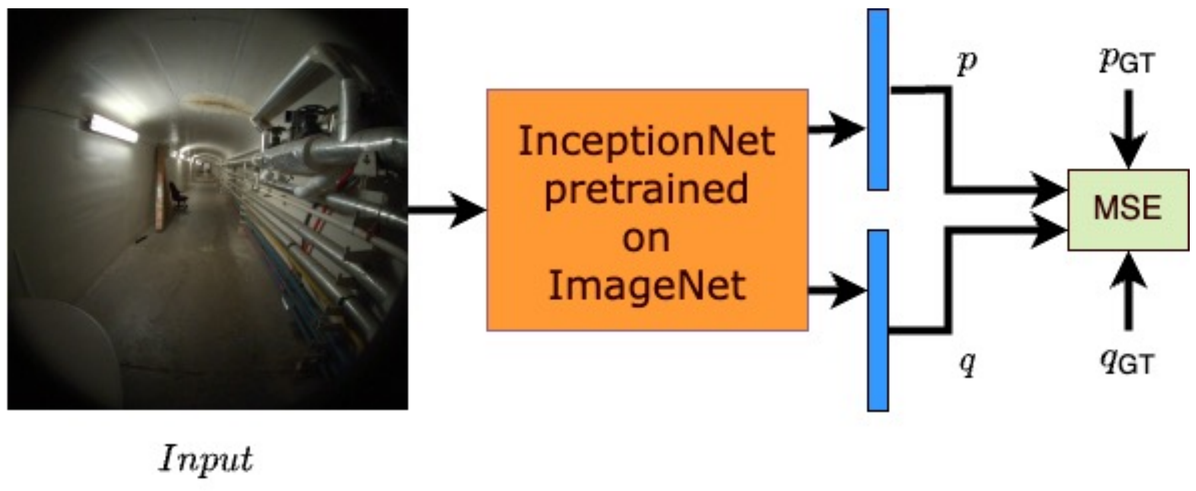
$$Magnetometer: m_x, m_y, m_z$$

Camera Capture:



20 mins trajectory in Smith Hall

6D Pose Estimation using RGB:



Model Architecture	King's College Dataset	FLOPs
InceptionNet pretrained on ImageNet Dataset	4.5 m; 9°	4.4 G
ResNet pretrained on Places dataset	4.3 m; 8.8°	7.6 G

Model Architecture	Smith Hall Dataset
InceptionNet pretrained on ImageNet Dataset	0.05 m; 10°

Deep Inertial Odometry:

Metric	IDOL Bldg. 1	IDOL Bldg. 2	Smith Hall
RMSE	14.9°	27.50°	17.2°

Orientation RMSE in Degrees

Metric	IDOL Bldg. 1	IDOL Bldg. 2	Smith Hall
ATE	5.67	10.33	3.9

Absolute Trajectory Error in Meters

Next Steps:

- Make algorithm location independent – Instead of estimating q using the Orientation Module, take orientation estimates from the device and use the Position Module as it is.
- Using 2 consecutive camera captures, get relative pose estimates and use these as the correction step in the Kalman Filter update, to estimate the final position and orientation.

References:

Kendall et al. "PoseNet: A Convolutional Network for Real-Time 6-DOF Camera Relocalization" | Sun, Scott et al. "IDOL: Inertial Deep Orientation-Estimation and Localization"