Introduction

- Self-driving vehicles need robust, computationally efficient visual-inertial SLAM for safe, real-time navigation.
- Modern visual-inertial SLAM systems require expensive sensors accompanied by extensive data.
- We lower the hardware and computational requirements by fundamentally improving the efficiency of information collection and processing from three aspects.

Active Sensing

- Exploration of the environment can increase confidence in map information during navigation.
- Instead of relying on sensors that passively collect data, we provide each camera with two more degrees of freedom via controllable pan-tilt units.
- This novel active sensing paradigm can provide autonomous agents with the means to maximize map coverage.

Active Sensing

• Monocular depth inference can be improved by considering camera intrinsics.
• Combining monocular depth inference and stereoscopic reconstruction could be a more resourceful alternative to depth sensors.

Online Back-End

• The block online EM SLAM only needs to estimate the recent trajectory by encoding the old information in the landmark estimates.
• Therefore, the block online EM SLAM facilitates a real-time visual-inertial system with an explicit map.
• In each block:
  \[ \hat{\lambda}_{k,0} = \hat{\lambda}_{k-1}, \]
  \[ Q_{k,n}(\lambda) = E_{\lambda_{k,n-1}} \left[ \log p(S_{Tk:T_{k+r_k}}, O_{Tk:T_{k+r_k}}; \lambda) | O_{Tk:T_{k+r_k}} \right], \]
  \[ \hat{\lambda}_{k,n} = \arg \max_{\lambda} Q_{k,n}(\lambda). \]