Adaptive Innovation Gating for Monocular Visual-Inertial State Estimation
Yuzhang Wu, Zheng Rong, Nathan Michael

Background
The quest to build smaller, more agile micro aerial vehicles has led to addressing cameras and IMUs as the primary sensors for state estimation. It is called visual-inertial state estimation. The minimum sensor suite only consists of a single camera and IMU.

Problem
State estimation with Visual Odometry cannot consistently achieve high performance due to:
- Features’ different properties
- Changing illumination conditions
- Various moving accelerated speed
- Combination of far and near objects

Solutions
Analyze exact factors affecting the accuracy of estimation results and the relationship between them
Increase the robustness of monocular visual-inertial state estimation using adaptive techniques.

Methods
Apparatus
- ARM computer
- Calibrated IMU
- Calibrated fish-eye monocular
Simulation
- Make synthetic datasets with diverse scene sizes
- Find relationship between feature depth and vehicle position estimated errors

Experiment
- Improve the algorithm
- Compare the robustness between original and adaptive programs

State Estimation Model
- Separately make models of IMU data and camera image to calculate the location of the feature
- Sliding windows:
  - Achieve baseline estimation
  - Decrease computing cost
  - Refine its solution from multi different observations

Innovation Gate
\[
\begin{align*}
\max & \left\{ h_0 - \lambda F \right\} \quad \text{subject to} \\
& \sum_{i=1}^{N} \left( \lambda e_i - h_i \right)^2 + \sum_{i=1}^{N} \left( \mathbf{w}_i - \mathbf{h}_i \right)^2 \leq \epsilon
\end{align*}
\]

Adaptation
- Improve the algorithm
- Compute the mean parallax of all features
- If the parallax of \( L \) feature is less than mean, then add it to the list of far features
- Unless the number of far feature is more than 30% of number of all features, then eliminate them

Comparison
- Evaluate both algorithms on an environment with both far and close features

Uncertainty
- Parallax \( \sum \frac{e_i}{h_i} \) > \( \epsilon \)
- For a definite depth, location uncertainty is inversely proportional with baseline.

Experiments
- Set 11 scenes with size from 10m to 60m, which means the depths are ranging
- Let the features distributed in margins of the environment.
- Analyze the situations in which the monocular state estimation will fail

References

Acknowledgments
Special thanks to Prof. Nathan, Zheng and John for tolerating my barrage of questions. Thanks too to the RISE program and NIST for giving me this unique opportunity.

Feedback: yuzhangw@ece.cmu.edu